## REMARKS

**Claim Status**

Claims 7, 21, 34 and 36 have been amended to recite that the microarray has a density of 1,000 or more regions per $cm^2$ or density of at least 1,000 regions/$cm^2$, and that the microarray comprises 1,000 or more regions. Support for the amendment of density of 1,000 or more regions per $cm^2$ or at least 1,000 regions/$cm^2$ is found on page 6, lines 3-5; page 10, lines 3-5 and page 12, lines 21-22 of the instant specification, wherein it is described that the microarray density is at least 1,000/$cm^2$. Support for the amendment of "1,000 or more regions" or "at least 1,000 regions" is found on page 32, lines 15-20 of the instant specification, which states that in a "general embodiment" of the invention, the microarray is created by arraying the DNA sequences in "1.0 $cm^2$". *See,* also page 42, lines 1-5.

In addition, claim 21 has also been amended to recite that the DNA sequences comprised in the microarray are polynucleotides of at least 50 subunits in length. Support for this amendment is found throughout the instant specification as well as in pending claims 7, 34 and 36.

Furthermore, claims 7-9, 21-23 and 34 have been amended to remove the term "about" from the phrases "about...or more" and "at least about".

Claim 16 has been amended to correct a typographic error.

Applicants respectfully submit that the foregoing amendments do not introduce any new material into the present application. With the present amendments, there are thirty-two claims pending, namely claims 7-19, 21-27 and 29-40.

**Priority**

The Examiner states that parent application Serial Nos. 08/514,875; 08/477,809; and 08/261,388 upon which priority is claimed do not provide support under 35 U.S.C. §112 for claims 14, 29, 35 and 38-39 of the present application. Applicants respectfully disagree for the following reasons.

With respect to instant claims 14 and 29, drawn to "covalently bound DNA", Applicants submit that such element is at least supported by the parent application Serial Nos. 08/477,809 and 08/514,875, wherein both specifications describe "the surface is a relatively hydrophilic, i.e., wettable surface, such as a surface having native, bound or covalently attached charged groups (emphasis added)." *See*, page 14, lines 17-20 of the parent application Serial No. 08/477,809 (the entire specification is shown in Exhibit A), and page 11, lines 28-31 of the parent application Serial No. 08/514,875 (the entire specification is downloaded from the Patent Office PAIR website and shown in Exhibit B). It is noted that DNA can be negatively charged. As such, the above citation indicates that charged groups, for example, DNA molecules, would be covalently bound to the surface of the substrate. Accordingly, claims 14 and 29 are at least entitled to the filing date of the parent application Serial No. 08/477,809, that is, June 7, 1995.

With respect to instant claim 35, drawn to "two fold change in the relative abundance", Applicants submit that such element is at least supported by the parent application Serial No. 08/514,875, wherein the specification describes "(t)he test average pattern is then compared with the control average pattern, to identify those test genes which show significantly, typically at least 2 fold and up to 100 fold or more, increase or decrease in gene expression level with respect to control levels for the same gene (emphasis added)." *See*, page 37, lines 21-26 of the parent

application Serial No. 08/514,875 (Exhibit B). Accordingly, claim 35 is at least entitled to the filing date of the parent application Serial No. 08/514,875, that is, August 14, 1995.

With respect to instant claims 38-39, drawn to "distinct gene sequences whose expression levels are specifically related to the differences between test cells relative to control cells", Applicants submit that this element is at least supported by the parent application Serial No. 08/514,875, wherein the specification describes that the subarray contains genes "whose gene expression levels are specifically related to the differences between test cells relative to control cells". *See*, page 4, lines 23-26 of the parent application Serial No. 08/514,875 (Exhibit B). Accordingly, claims 38-39 are at least entitled to the filing date of the parent application Serial No. 08/514,875, that is, August 14, 1995.

**Claim Objections**

Claim 16 is objected to because "polycationic" is misspelled. In response, Applicants have amended claim 16 to correct this typographic error as suggested by the Examiner.

**Rejection under 35 U.S.C. § 112, First Paragraph (Written Description)**

Claims 7-19, 21-27, 29-40 are rejected under 35 U.S.C. § 112, first paragraph, as allegedly failing to comply with the written description requirement. Applicants respectfully traverse this rejection.

Claims 7, 21, 34 and 36 have been amended to recite that the microarray has a density of 1,000 or more regions per $cm^2$ or density of at least 1,000 regions/$cm^2$, and that the microarray comprises 1,000 or more regions. As discussed above, direct written support for the presently

amended claims is evident from the instant specification. With the density range being 1,000 or more regions per $cm^2$ or at least 1,000 regions/$cm^2$ as presently amended and the size of the microarray being 1.0 $cm^2$ as described in the instant specification, Applicants submit that the instant microarray obviously comprises 1,000 or more regions.

In view of the above remarks, Applicants believe that the claims as presently amended fully comply with the written description requirement and as such, respectfully request the withdrawn of the present rejection under 35 U.S.C. §112, first paragraph.

**Rejection under 35 U.S.C. § 112, Second Paragraph**

Claims 7-19, 21-27 and 29-40 are rejected under 35 U.S.C. §112, second paragraph, as allegedly being indefinite. In response, Applicants have amended claims 7-9, 21-23 and 34 by removing the term "about" from the phrases "about...or more" and "at least about" therein. Consequently, the instant claims as presently amended are definite and the rejection under 35 U.S.C. § 112, second paragraph, should be withdrawn.

**Rejection under 35 U.S.C. § 102**

Claims 7-19, 21-27 and 29-40 are rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Fodor et al. (U.S. Pat. No. 6,610,482, referred to as "the '482 patent" hereafter). Applicants respectfully traverse this rejection.

In asserting this rejection, it appears that the Examiner considers December 6, 1990 to be the effective filing date of the '482 patent. Applicants contend that the '482 patent is not entitled to the benefit of that date.

The '482 patent claims priority of an earlier application Serial No. 08/670,118, filed on June 25, 1996, subsequently issued as U.S. Patent No. 5,800,992 (*see*, Exhibit C.) However, it is believed that the 08/670,118 application upon which the '482 patent claims priority does not provide any support for the element of "greater than 50 monomers in length" as recited in claims 40 and 56 of the '482 patent. Applicants present the following detailed reasoning.

The 08/670,118 application discloses an *in situ* synthesis procedure for production of polymer arrays (polypeptide or oligonucleotide arrays), which procedure was called the Very Large Scale Immobilized Polymer Synthesis (VLSIPS). The 08/670,118 application further describes that VLSIPS technology certainly allows production of ten nucleotide oligomers on a solid phase (*see,* col. 19, lines 36-38 of the issued Patent No. 5,800,992, Exhibit C) and that the length of oligonucleotides used will be selected on criteria determined to some extent by practical limits (*see* col. 20, lines 62 – col. 21, line 7 of the issued Patent No. 5,800,992, Exhibit C). The passage goes on as cited below:

> For example, if probes are made as oligonucleotides, there will be 65,536 possible eight nucleotide sequences. If a nine subunit oligonucleotide is selected, there are 262,144 possible permeations of sequences. If a ten-mer oligonucleotide is selected, there are 1,048,576 possible permeations of sequences. As the number gets larger, the required number of positionally defined subunits necessary to saturate the possibilities also increases. With respect to hybridization conditions, the length of the matching necessary to confer stability of the conditions selected can be compensated for.

Col. 22, lines 6-12 of the 5,800,992 (Exhibit C) patent further describes following limitations of the VLSIPS technology in sequencing:

> As the length of the oligomer increases the number of different probes which must be synthesized also increases at a rate of a factor of 4 for every additional nucleotide. Eventually the size of the matrix and the limitations in the resolution of regions in the matrix will reach the point where an increase in number of probes becomes disadvantageous.

The above descriptions of the 08/670,118 application indicate that if a 50-mer oligonucleotide were selected, there would be $1,048,576^5$ possible permeation sequences. Such an astronomic number of probes that must be synthesized on a matrix would certainly exceed the size and resolution limits of the matrix provided by the 08/670,118 application. In fact, the 08/670,118 application does not mention anything about producing an array comprising polymers having greater than 50 monomers in length.

That is, by the time of the filing of the application Serial No. 08/670,118, i.e., June 25, 1996, the Fodor group was working on VLSIPS technology, which could not have produced a microarray of polynucleotides each having more than 50 monomeric units. As such, Applicants believe that the later patent of Fodor's, i.e., the '482 patent cited by the Examiner in the instant case, should not have been entitled the priority of the 08/670,118 application at least when considering the element of "greater than 50 monomers in length" as recited in claims 40 and 56 of the '482 patent. The effective filing date for claims 40 and 56 of the '482 patent should be later than the filing date of the 08/670,118 application, that is, later than June 25, 1996.

Applicants further submit that, when considering the element of "at least 50 subunits" as recited in instant claims 7, 34 and 36, the present application is entitled to at least the filing date of parent application Serial No. 08/477,809, that is, June 7, 1995. It is noted that such element is at least supported by the parent application Serial No. 08/477,809, wherein the specification describes "(e)ach distinct biopolymer (i) is disposed at a separate, defined position in said array, (ii) has a length of at least 50 subunits" (*see*, page 7, lines 26-28, Exhibit A), "the biopolymers are polynucleotides having lengths of at least about 50 bp" (*see*, page 26, lines 21-23, Exhibit A),

and "the polynucleotides have lengths of at least about 50 bp" (*see*, page 29, lines 5-6, Exhibit A).

As such, claims 7, 34 and 36 of the present application is at least entitled the priority date of June 7, 1995, which date is well before the priority date for claims 40 and 56 of the '482 patent (i.e., later than June 25, 1996). As such, the '482 patent is not a valid prior art reference for claims 7, 34 and 36 of the present application. Consequently, the rejection of claims 7-19 and 34-39 under 35 U.S.C §102(e) as allegedly being anticipated by the '482 patent should be withdrawn.

With respect to claims 21-33 and 40 of the present application, Applicants have amended claim 21 to recite that the DNA sequences in the microarray are polynucleotides of at least 50 subunits in length. Same as claims 7-19 and 34-39, claim 21 as presently amended and its dependent claims (i.e., claims 21-33 and 40) would have an earlier priority date than that of the '482 patent when considering the element of "at least 50 subunits".

In view of the above remarks, Applicants believe that the '482 patent is not qualified to be a prior art reference for the present novelty rejection as the '482 patent has a later priority date than that of the present application when considering the particular element of "at least 50 subunits". Therefore, the rejection of the present claims under 35 U.S.C. §102 in view of the '482 patent should be withdrawn.

Claims 7-15, 17-19, 21-27, 29-35 and 38-40 are rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Winkler et al. (U.S. Pat. No. 6,677,195). Applicants respectfully traverse this rejection.

Winkler et al. discloses a method and device for forming large arrays of polymers on a substrate. In particular, Winkler et al. discloses a monomer-by-monomer synthesis process, wherein the deposition of monomers is carried out by using a deposition device to effect the formation of a polymer in a pre-defined region. That is, the arrays of polymers are formed through an in situ process. Although it mentions that "the process is repeated to provide polymers with as few as two monomers, although the process may be readily adapted to form polymers having 3, 4, 5, 6, 10, 15, 20, 30, 40, 50, 75, 100 or more monomers therein" (*see*, col. 17, lines 53-57), Winkler et al. does not provide any enabling data that arrays of "polymers having 3, 4, 5, 6, 10, 15, 20, 30, 40, 50, 75, 100 or more monomers" could actually be produced using their monomer-by-monomer synthesis process. In fact, the state of the art at the time of the filing of Winkler et al. (i.e., Nov. 1992) suggests that numerous inherent drawbacks of in situ synthesis including low coupling efficiency and premature truncation of a polymer limit the length of the synthesized polymers and their homogeneity. Mere mentioning of forming "polymers having 3, 4, 5, 6, 10, 15, 20, 30, 40, 50, 75, 100 or more monomers therein" does not suggest that such idea was reduced to practice at the time of the filing of Winkler et al.

In fact, Applicants submit that Winkler et al. and the above-discussed 08/670,118 application share a common inventor (i.e., Stephen P.A. Fodor) and that Winkler et al. predated the filing date of the 08/670,118 application (i.e., June 25, 1996). As recognized in the 08/670,118 application, which apparently describes the later work of Fodor's group, in situ synthesis has such limitations in sequencing that producing a 50-mer oligonucleotide would seem to be impossible due to astronomic number of probes that must be synthesized on a then-available matrix. Taking into consideration the state of the art at the time of its filing as well as

the fact that it lacks any enabling data and factual support, Applicants contend that Winkler et al. would not have been able to produce a microarray of polymers having greater than 50 monomers in length. As such, Winkler et al. could not anticipate the present invention as claimed. Therefore, the present novelty rejection in view of Winkler et al. should be withdrawn.

* * *

This paper is filed concurrently with a petition for a three-month extension of time. The Commissioner is authorized to deduct the extension fee ($510) from Howrey LLP Deposit Account No. 08-3038/12665.0009.CNUS01. Should any additional fees be required for any reason relating to this document, the Commissioner is authorized to deduct such fees from the same Deposit Account.

Respectfully submitted,

J. Wendy Davis, Ph.D.
Reg. No. 46,393

Agent for Assignee
THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY

Customer No. 23369

HOWREY LLP
1111 Louisiana, 25th floor
Houston, TX 77002
(713) 787-1512 (Direct)

Date: September 14, 2006

EXPRESS MAIL MAILING LABEL

NUMBER: EV 691741584 US
DATE OF DEPOSIT: September 14, 2006

I hereby certify that this paper or fee is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. §1.10 on the date indicated above and is addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria VA, 22313-1450.

Paula S. Linkhart

Paula S. Linkhart

# Specification as originally filed
# Serial No. 08/477,809

Attorney Docket No.: 8600-0138.30 Patent

**Transmittal of Utility Patent Application for Filing**

*Certification Under 37 C.F.R. §1.10 (if applicable)*

TB 221 728 914 US
"Express Mail" Label Number

7 June 1995
Date of Deposit

I hereby certify that this Transmittal letter, enclosed application, and any other documents referred to as enclosed herein are being deposited in an envelope with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231

Peter J. Dehlinger
(Print Name of Person Mailing Application)

Peter J. Dehlinger
(Signature of Person Mailing Application)

# METHOD AND APPARATUS FOR FABRICATING MICROARRAYS OF BIOLOGICAL SAMPLES

## Field of the Invention

This invention relates to a method and apparatus for fabricating microarrays of biological samples for large scale screening assays, such as arrays of DNA samples to be used in DNA hybridization assays for genetic research and diagnostic applications.

## References


Abouzied, *et al.*, *Journal of AOAC International* 77(2):495-500 (1994).

Bohlander, *et al.*, *Genomics* 13:1322-1324 (1992).

Drmanac, *et al.*, *Science* 260:1649-1652 (1993).

Fodor, *et al.*, *Science* 251:767-773 (1991).

Khrapko, *et al.*, *DNA Sequence* 1:375-388 (1991).

Kuriyama, *et al.*, AN ISFET BIOSENSOR, APPLIED BIOSENSORS (Donald Wise, Ed.), Butterworths, pp. 93-114 (1989).

Lehrach, *et al.*, HYBRIDIZATION FINGERPRINTING IN GENOME MAPPING AND SEQUENCING, GENOME ANALYSIS, VOL 1 (Davies and Tilgham, Eds.), Cold Spring Harbor Press, pp. 39-81 (1990).

Maniatis, *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press (1989).

Nelson, *et al.*, *Nature Genetics* 4:11-18 (1993).

Pirrung, *et al.*, U.S. Patent No. 5,143,854 (1992).

Riles, *et al.*, *Genetics* 134:81-150 (1993).

Schena, M. *et al.*, *Proc. Nat. Acad. Sci. USA* 89:3894-3898 (1992).

Southern, *et al.*, *Genomics* 13:1008-1017 (1992).

**Background of the Invention**

A variety of methods are currently available for making arrays of biological macromolecules, such as arrays of nucleic acid molecules or proteins. One method for making ordered arrays of DNA on a porous membrane is a "dot blot" approach. In this method, a vacuum manifold transfers a plurality, *e.g.*, 96, aqueous samples of DNA from 3 millimeter diameter wells to a porous membrane. A common variant of this procedure is a "slot-blot" method in which the wells have highly-elongated oval shapes.

The DNA is immobilized on the porous membrane by baking the membrane or exposing it to UV radiation. This is a manual procedure practical for making one array at a time and usually limited to 96 samples per array. "Dot-blot" procedures are therefore inadequate for applications in which many thousand samples must be determined.

A more efficient technique employed for making ordered arrays of genomic fragments uses an array of pins dipped into the wells, *e.g.*, the 96 wells of a microtitre plate, for transferring an array of samples to a substrate, such as a porous membrane. One array includes pins that are designed to spot a membrane in a staggered fashion, for creating an array of 9216 spots in a 22 × 22 cm area (Lehrach, *et al.*, 1990). A limitation with this approach is that the volume of DNA spotted in each pixel of each array is highly variable.

In addition, the number of arrays that can be made with each dipping is usually quite small.

An alternate method of creating ordered arrays of nucleic acid sequences is described by Pirrung, *et al.* (1992), and also by Fodor, *et al.* (1991). The method involves synthesizing different nucleic acid sequences at different discrete regions of a support. This method employs elaborate synthetic schemes, and is generally limited to relatively short nucleic acid sample, *e.g.*, less than 20 bases. A related method has been described by Southern, *et al.* (1992).

Khrapko, *et al.* (1991) describes a method of making an oligonucleotide matrix by spotting DNA onto a thin layer of polyacrylamide. The spotting is done manually with a micropipette.

None of the methods or devices described in the prior art are designed for mass fabrication of microarrays characterized by (i) a large number of micro-sized assay regions separated by a distance of 50-200 microns or less, and (ii) a well-defined amount, typically in the picomole range, of analyte associated with each region of the array.

Furthermore, current technology is directed at performing such assays one at a time to a single array of DNA molecules. For example, the most common method for performing DNA hybridizations to arrays spotted onto porous membrane involves sealing the membrane in a plastic bag (Maniatas, *et al.*, 1989) or a rotating glass cylinder (Robbins Scientific) with the labeled hybridization probe inside the sealed chamber. For arrays made on non-porous surfaces, such as a microscope slide, each array is incubated with the labeled hybridization probe sealed under a coverslip. These techniques require a separate sealed chamber for

each array which makes the screening and handling of many such arrays inconvenient and time intensive.

Abouzied, *et al.* (1994) describes a method of printing horizontal lines of antibodies on a nitrocellulose membrane and separating regions of the membrane with vertical stripes of a hydrophobic material. Each vertical stripe is then reacted with a different antigen and the reaction between the immobilized antibody and an antigen is detected using a standard ELISA colorimetric technique. Abouzied's technique makes it possible to screen many one-dimensional arrays simultaneously on a single sheet of nitrocellulose. Abouzied makes the nitrocellulose somewhat hydrophobic using a line drawn with PAP Pen (Research Products International). However Abouzied does not describe a technology that is capable of completely sealing the pores of the nitrocellulose. The pores of the nitrocellulose are still physically open and so the assay reagents can leak through the hydrophobic barrier during extended high temperature incubations or in the presence of detergents which makes the Abouzied technique unacceptable for DNA hybridization assays.

Porous membranes with printed patterns of hydrophilic/hydrophobic regions exist for applications such as ordered arrays of bacteria colonies. QA Life Sciences (San Diego CA) makes such a membrane with a grid pattern printed on it. However, this membrane has the same disadvantage as the Abouzied technique since reagents can still flow between the gridded arrays making them unusable for separate DNA hybridization assays.

Pall Corporation make a 96-well plate with a porous filter heat sealed to the bottom of the plate. These plates are capable of containing different

reagents in each well without cross-contamination. However, each well is intended to hold only one target element whereas the invention described here makes a microarray of many biomolecules in each subdivided region of the solid support. Furthermore, the 96 well plates are at least 1 cm thick and prevent the use of the device for many colorimetric, fluorescent and radioactive detection formats which require that the membrane lie flat against the detection surface. The invention described here requires no further processing after the assay step since the barriers elements are shallow and do not interfere with the detection step thereby greatly increasing convenience.

Hyseq Corporation has described a method of making an "array of arrays" on a non-porous solid support for use with their sequencing by hybridization technique. The method described by Hyseq involves modifying the chemistry of the solid support material to form a hydrophobic grid pattern where each subdivided region contains a microarray of biomolecules. Hyseq's flat hydrophobic pattern does not make use of physical blocking as an additional means of preventing cross contamination.

## Summary of the Invention

The invention includes, in one aspect, a method of forming a microarray of analyte-assay regions on a solid support, where each region in the array has a known amount of a selected, analyte-specific reagent. The method involves first loading a solution of a selected analyte-specific reagent in a reagent-dispensing device having an elongate capillary channel (i) formed by spaced-apart, coextensive elongate members, (ii) adapted to hold a quantity of the reagent solution and (iii) having a tip region at which aqueous

solution in the channel forms a meniscus. The channel is preferably formed by a pair of spaced-apart tapered elements.

The tip of the dispensing device is tapped against a solid support at a defined position on the support surface with an impulse effective to break the meniscus in the capillary channel deposit a selected volume of solution on the surface, preferably a selected volume in the range 0.01 to 100 nl. The two steps are repeated until the desired array is formed.

The method may be practiced in forming a plurality of such arrays, where the solution-depositing step is are applied to a selected position on each of a plurality of solid supports at each repeat cycle.

The dispensing device may be loaded with a new solution, by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and (iii) dipping the capillary channel into the new reagent solution.

Also included in the invention is an automated apparatus for forming a microarray of analyte-assay regions on a plurality of solid supports, where each region in the array has a known amount of a selected, analyte-specific reagent. The apparatus has a holder for holding, at known positions, a plurality of planar supports, and a reagent dispensing device of the type described above.

The apparatus further includes positioning structure for positioning the dispensing device at a selected array position with respect to a support in said holder, and dispensing structure for moving the dispensing device into tapping engagement against a support with a selected impulse effective to deposit a

selected volume on the support, *e.g.*, a selected volume in the volume range 0.01 to 100 nl.

The positioning and dispensing structures are controlled by a control unit in the apparatus. The unit operates to (i) place the dispensing device at a loading station, (ii) move the capillary channel in the device into a selected reagent at the loading station, to load the dispensing device with the reagent, and (iii) dispense the reagent at a defined array position on each of the supports on said holder. The unit may further operate, at the end of a dispensing cycle, to wash the dispensing device by (i) placing the dispensing device at a washing station, (ii) moving the capillary channel in the device into a wash fluid, to load the dispensing device with the fluid, and (iii) remove the wash fluid prior to loading the dispensing device with a fresh selected reagent.

The dispensing device in the apparatus may be one of a plurality of such devices which are carried on the arm for dispensing different analyte assay reagents at selected spaced array positions.

In another aspect, the invention includes a substrate with a surface having a microarray of at least $10^3$ distinct polynucleotide or polypeptide biopolymers in a surface area of less than about 1 $cm^2$. Each distinct biopolymer (i) is disposed at a separate, defined position in said array, (ii) has a length of at least 50 subunits, and (iii) is present in a defined amount between about 0.1 femtomoles and 100 nanomoles.

In one embodiment, the surface is glass slide surface coated with a polycationic polymer, such as polylysine, and the biopolymers are polynucleotides. In another embodiment, the substrate has a water-impermeable backing, a water-permeable film formed on

the backing, and a grid formed on the film. The grid is composed of intersecting water-impervious grid elements extending from said backing to positions raised above the surface of said film, and partitions the film into a plurality of water-impervious cells. A biopolymer array is formed within each well.

More generally, there is provided a substrate for use in detecting binding of labeled polynucleotides to one or more of a plurality different-sequence, immobilized polynucleotides. The substrate includes, in one aspect, a glass support, a coating of a polycationic polymer, such as polylysine, on said surface of the support, and an array of distinct polynucleotides electrostatically bound non-covalently to said coating, where each distinct biopolymer is disposed at a separate, defined position in a surface array of polynucleotides.

In another aspect, the substrate includes a water-impermeable backing, a water-permeable film formed on the backing, and a grid formed on the film, where the grid is composed of intersecting water-impervious grid elements extending from the backing to positions raised above the surface of the film, forming a plurality of cells. A biopolymer array is formed within each cell.

Also forming part of the invention is a method of detecting differential expression of each of a plurality of genes in a first cell type, with respect to expression of the same genes in a second cell type. In practicing the method, there is first produced fluorescent-labeled cDNA's from mRNA's isolated from the two cells types, where the cDNA'S from the first and second cells are labeled with first and second different fluorescent reporters.

A mixture of the labeled cDNA's from the two cell types is added to an array of polynucleotides

representing a plurality of known genes derived from the two cell types, under conditions that result in hybridization of the cDNA's to complementary-sequence polynucleotides in the array. The array is then examined by fluorescence under fluorescence excitation conditions in which (i) polynucleotides in the array that are hybridized predominantly to cDNA's derived from one of the first and second cell types give a distinct first or second fluorescence emission color, respectively, and (ii) polynucleotides in the array that are hybridized to substantially equal numbers of cDNA's derived from the first and second cell types give a distinct combined fluorescence emission color, respectively. The relative expression of known genes in the two cell types can then be determined by the observed fluorescence emission color of each spot.

These and other objects and features of the invention will become more fully apparent when the following detailed description of the invention is read in conjunction with the accompanying figures.

## Brief Description of the Drawings

Fig. 1 is a side view of a reagent-dispensing device having a open-capillary dispensing head constructed for use in one embodiment of the invention;

Figs. 2A-2C illustrate steps in the delivery of a fixed-volume bead on a hydrophobic surface employing the dispensing head from Fig. 1, in accordance with one embodiment of the method of the invention;

Fig. 3 shows a portion of a two-dimensional array of analyte-assay regions constructed according to the method of the invention;

Fig. 4 is a planar view showing components of an automated apparatus for forming arrays in accordance with the invention.

Fig. 5 shows a fluorescent image of an actual 20 × 20 array of 400 fluorescently-labeled DNA samples immobilized on a poly-l-lysine coated slide, where the total area covered by the 400 element array is 16 square millimeters;

Fig. 6 is a fluorescent image of a 1.8 cm × 1.8 cm microarray containing lambda clones with yeast inserts, the fluorescent signal arising from the hybridization to the array with approximately half the yeast genome labeled with a green fluorophore and the other half with a red fluorophore;

Fig. 7 shows the translation of the hybridization image of Fig. 6 into a karyotype of the yeast genome, where the elements of Fig.-6 microarray contain yeast DNA sequences that have been previously physically mapped in the yeast genome;

Fig. 8 show a fluorescent image of a 0.5 cm × 0.5 cm microarray of 24 cDNA clones, where the microarray was hybridized simultaneously with total cDNA from wild type *Arabidopsis* plant labeled with a green fluorophore and total cDNA from a transgenic *Arabidopsis* plant labeled with a red fluorophore, and the arrow points to the cDNA clone representing the gene introduced into the transgenic *Arabidopsis* plant;

Fig. 9 shows a plan view of substrate having an array of cells formed by barrier elements in the form of a grid;

Fig. 10 shows an enlarged plan view of one of the cells in the substrate in Fig. 9, showing an array of polynucleotide regions in the cell;

Fig. 11 is an enlarged sectional view of the substrate in Fig. 9, taken along a section line in that figure; and

Fig. 12 is a scanned image of a 3 cm × 3 cm nitrocellulose solid support containing four identical

arrays of M13 clones in each of four quadrants, where each quadrant was hybridized simultaneously to a different oligonucleotide using an open face hybridization method.

## Detailed Description of the Invention

### I. Definitions

Unless indicated otherwise, the terms defined below have the following meanings:

"Ligand" refers to one member of a ligand/anti-ligand binding pair. The ligand may be, for example, one of the nucleic acid strands in a complementary, hybridized nucleic acid duplex binding pair; an effector molecule in an effector/receptor binding pair; or an antigen in an antigen/antibody or antigen/antibody fragment binding pair.

"Antiligand" refers to the opposite member of a ligand/anti-ligand binding pair. The antiligand may be the other of the nucleic acid strands in a complementary, hybridized nucleic acid duplex binding pair; the receptor molecule in an effector/receptor binding pair; or an antibody or antibody fragment molecule in antigen/antibody or antigen/antibody fragment binding pair, respectively.

"Analyte" or "analyte molecule" refers to a molecule, typically a macromolecule, such as a polynucleotide or polypeptide, whose presence, amount, and/or identity are to be determined. The analyte is one member of a ligand/anti-ligand pair.

"Analyte-specific assay reagent" refers to a molecule effective to bind specifically to an analyte molecule. The reagent is the opposite member of a ligand/anti-ligand binding pair.

An "array of regions on a solid support" is a linear or two-dimensional array of preferably discrete

regions, each having a finite area, formed on the surface of a solid support.

A "microarray" is an array of regions having a density of discrete regions of at least about 100/cm$^2$, and preferably at least about 1000/cm$^2$. The regions in a microarray have typical dimensions, *e.g.*, diameters, in the range of between about 10-250 $\mu$m, and are separated from other regions in the array by about the same distance.

A support surface is "hydrophobic" if a aqueous-medium droplet applied to the surface does not spread out substantially beyond the area size of the applied droplet. That is, the surface acts to prevent spreading of the droplet applied to the surface by hydrophobic interaction with the droplet.

A "meniscus" means a concave or convex surface that forms on the bottom of a liquid in a channel as a result of the surface tension of the liquid.

"Distinct biopolymers", as applied to the biopolymers forming a microarray, means an array member which is distinct from other array members on the basis of a different biopolymer sequence, and/or different concentrations of the same or distinct biopolymers, and/or different mixtures of distinct or different-concentration biopolymers. Thus an array of "distinct polynucleotides" means an array containing, as its members, (i) distinct polynucleotides, which may have a defined amount in each member, (ii) different, graded concentrations of given-sequence polynucleotides, and/or (iii) different-composition mixtures of two or more distinct polynucleotides.

"Cell type" means a cell from a given source, *e.g.*, a tissue, or organ, or a cell in a given state of

differentiation, or a cell associated with a given pathology or genetic makeup.

II. Method of Microarray Formation

This section describes a method of forming a microarray of analyte-assay regions on a solid support or substrate, where each region in the array has a known amount of a selected, analyte-specific reagent.

Fig. 1 illustrates, in a partially schematic view, a reagent-dispensing device 10 useful in practicing the method. The device generally includes a reagent dispenser 12 having an elongate open capillary channel 14 adapted to hold a quantity of the reagent solution, such as indicated at 16, as will be described below. The capillary channel is formed by a pair of spaced-apart, coextensive, elongate members 12a, 12b which are tapered toward one another and converge at a tip or tip region 18 at the lower end of the channel. More generally, the open channel is formed by at least two elongate, spaced-apart members adapted to hold a quantity of reagent solutions and having a tip region at which aqueous solution in the channel forms a meniscus, such as the concave meniscus illustrated at 20 in Fig. 2A. The advantages of the open channel construction of the dispenser are discussed below.

With continued reference to Fig. 1, the dispenser device also includes structure for moving the dispenser rapidly toward and away from a support surface, for effecting deposition of a known amount of solution in the dispenser on a support, as will be described below with reference to Figs. 2A-2C. In the embodiment shown, this structure includes a solenoid 22 which is activatable to draw a solenoid piston 24 rapidly downwardly, then release the piston, *e.g.*, under spring bias, to a normal, raised position, as shown. The

dispenser is carried on the piston by a connecting member 26, as shown. The just-described moving structure is also referred to herein as dispensing means for moving the dispenser into engagement with a solid support, for dispensing a known volume of fluid on the support.

The dispensing device just described is carried on an arm 28 that may be moved either linearly or in an x-y plane to position the dispenser at a selected deposition position, as will be described.

Figs. 2A-2C illustrate the method of depositing a known amount of reagent solution in the just-described dispenser on the surface of a solid support, such as the support indicated at 30. The support is a polymer, glass, or other solid-material support having a surface indicated at 31.

In one general embodiment, the surface is a relatively hydrophilic, *i.e.*, wettable surface, such as a surface having native, bound or covalently attached charged groups. On such surface described below is a glass surface having an absorbed layer of a polycationic polymer, such as poly-l-lysine.

In another embodiment, the surface has or is formed to have a relatively hydrophobic character, *i.e.*, one that causes aqueous medium deposited on the surface to bead. A variety of known hydrophobic polymers, such as polystyrene, polypropylene, or polyethylene have desired hydrophobic properties, as do glass and a variety of lubricant or other hydrophobic films that may be applied to the support surface.

Initially, the dispenser is loaded with a selected analyte-specific reagent solution, such as by dipping the dispenser tip, after washing, into a solution of the reagent, and allowing filling by capillary flow into the dispenser channel. The dispenser is now moved

to a selected position with respect to a support surface, placing the dispenser tip directly above the support-surface position at which the reagent is to be deposited. This movement takes place with the dispenser tip in its raised position, as seen in Fig. 2A, where the tip is typically at least several 1-5 mm above the surface of the substrate.

With the dispenser so positioned, solenoid 22 is now activated to cause the dispenser tip to move rapidly toward and away from the substrate surface, making momentary contact with the surface, in effect, tapping the tip of the dispenser against the support surface. The tapping movement of the tip against the surface acts to break the liquid meniscus in the tip channel, bringing the liquid in the tip into contact with the support surface. This, in turn, produces a flowing of the liquid into the capillary space between the tip and the surface, acting to draw liquid out of the dispenser channel, as seen in Fig. 2B.

Fig. 2C shows flow of fluid from the tip onto the support surface, which in this case is a hydrophobic surface. The figure illustrates that liquid continues to flow from the dispenser onto the support surface until it forms a liquid bead 32. At a given bead size, *i.e.*, volume, the tendency of liquid to flow onto the surface will be balanced by the hydrophobic surface interaction of the bead with the support surface, which acts to limit the total bead area on the surface, and by the surface tension of the droplet, which tends toward a given bead curvature. At this point, a given bead volume will have formed, and continued contact of the dispenser tip with the bead, as the dispenser tip is being withdrawn, will have little or no effect on bead volume.

For liquid-dispensing on a more hydrophilic surface, the liquid will have less of a tendency to bead, and the dispensed volume will be more sensitive to the total dwell time of the dispenser tip in the immediate vicinity of the support surface, *e.g.*, the positions illustrated in Figs. 2B and 2C.

The desired deposition volume, *i.e.*, bead volume, formed by this method is preferably in the range 2 pl (picoliters) to 2 nl (nanoliters), although volumes as high as 100 nl or more may be dispensed. It will be appreciated that the selected dispensed volume will depend on (i) the "footprint" of the dispenser tip, *i.e.*, the size of the area spanned by the tip, (ii) the hydrophobicity of the support surface, and (iii) the time of contact with and rate of withdrawal of the tip from the support surface. In addition, bead size may be reduced by increasing the viscosity of the medium, effectively reducing the flow time of liquid from the dispenser onto the support surface. The drop size may be further constrained by depositing the drop in a hydrophilic region surrounded by a hydrophobic grid pattern on the support surface.

In a typical embodiment, the dispenser tip is tapped rapidly against the support surface, with a total residence time in contact with the support of less than about 1 msec, and a rate of upward travel from the surface of about 10 cm/sec.

Assuming that the bead that forms on contact with the surface is a hemispherical bead, with a diameter approximately equal to the width of the dispenser tip, as shown in Fig. 2C, the volume of the bead formed in relation to dispenser tip width (d) is given in Table 1 below. As seen, the volume of the bead ranges between 2 pl to 2 nl as the width size is increased from about 20 to 200 $\mu$m.

**Table 1**

| d | Volume (nl) |
|---|---|
| 20 $\mu$m | $2 \times 10^{-3}$ |
| 50 $\mu$m | $3.1 \times 10^{-2}$ |
| 100 $\mu$m | $2.5 \times 10^{-1}$ |
| 200 $\mu$m | 2 |

At a given tip size, bead volume can be reduced in a controlled fashion by increasing surface hydrophobicity, reducing time of contact of the tip with the surface, increasing rate of movement of the tip away from the surface, and/or increasing the viscosity of the medium. Once these parameters are fixed, a selected deposition volume in the desired pl to nl range can be achieved in a repeatable fashion.

After depositing a bead at one selected location on a support, the tip is typically moved to a corresponding position on a second support, a droplet is deposited at that position, and this process is repeated until a liquid droplet of the reagent has been deposited at a selected position on each of a plurality of supports.

The tip is then washed to remove the reagent liquid, filled with another reagent liquid and this reagent is now deposited at each another array position on each of the supports. In one embodiment, the tip is washed and refilled by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and (iii) dipping the capillary channel into the new reagent solution.

From the foregoing, it will be appreciated that the tweezers-like, open-capillary dispenser tip

provides the advantages that (i) the open channel of the tip facilitates rapid, efficient washing and drying before reloading the tip with a new reagent, (ii) passive capillary action can load the sample directly from a standard microwell plate while retaining sufficient sample in the open capillary reservoir for the printing of numerous arrays, (iii) open capillaries are less prone to clogging than closed capillaries, and (iv) open capillaries do not require a perfectly faced bottom surface for fluid delivery.

A portion of a microarray 36 formed on the surface 38 of a solid support 40 in accordance with the method just described is shown in Fig. 3. The array is formed of a plurality of analyte-specific reagent regions, such as regions 42, where each region may include a different analyte-specific reagent. As indicated above, the diameter of each region is preferably between about 20-200 $\mu$m. The spacing between each region and its closest (non-diagonal) neighbor, measured from center-to-center (indicated at 44), is preferably in the range of about 20-400 $\mu$m. Thus, for example, an array having a center-to-center spacing of about 250 $\mu$m contains about 40 regions/cm or 1,600 regions/cm$^2$. After formation of the array, the support is treated to evaporate the liquid of the droplet forming each region, to leave a desired array of dried, relatively flat regions. This drying may be done by heating or under vacuum.

In some cases, it is desired to first rehydrate the droplets containing the analyte reagents to allow for more time for adsorption to the solid support. It is also possible to spot out the analyte reagents in a humid environment so that droplets do not dry until the arraying operation is complete.

III. Automated Apparatus for Forming Arrays

In another aspect, the invention includes an automated apparatus for forming an array of analyte-assay regions on a solid support, where each region in the array has a known amount of a selected, analyte-specific reagent.

The apparatus is shown in planar, and partially schematic view in Fig. 4. A dispenser device 72 in the apparatus has the basic construction described above with respect to Fig. 1, and includes a dispenser 74 having an open-capillary channel terminating at a tip, substantially as shown in Figs. 1 and 2A-2C.

The dispenser is mounted in the device for movement toward and away from a dispensing position at which the tip of the dispenser taps a support surface, to dispense a selected volume of reagent solution, as described above. This movement is effected by a solenoid 76 as described above. Solenoid 76 is under the control of a control unit 77 whose operation will be described below. The solenoid is also referred to herein as dispensing means for moving the device into tapping engagement with a support, when the device is positioned at a defined array position with respect to that support.

The dispenser device is carried on an arm 74 which is threadedly mounted on a worm screw 80 driven (rotated) in a desired direction by a stepper motor 82 also under the control of unit 77. At its left end in the figure screw 80 is carried in a sleeve 84 for rotation about the screw axis. At its other end, the screw is mounted to the drive shaft of the stepper motor, which in turn is carried on a sleeve 86. The dispenser device, worm screw, the two sleeves mounting the worm screw, and the stepper motor used in moving the device in the "x" (horizontal) direction in the

figure form what is referred to here collectively as a displacement assembly 86.

The displacement assembly is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an x axis in the figure. In one mode, the assembly functions to move the dispenser in x-axis increments having a selected distance in the range 5-25 μm. In another mode, the dispenser unit may be moved in precise x-axis increments of several microns or more, for positioning the dispenser at associated positions on adjacent supports, as will be described below.

The displacement assembly, in turn, is mounted for movement in the "y" (vertical) axis of the figure, for positioning the dispenser at a selected y axis position. The structure mounting the assembly includes a fixed rod 88 mounted rigidly between a pair of frame bars 90, 92, and a worm screw 94 mounted for rotation between a pair of frame bars 96, 98. The worm screw is driven (rotated) by a stepper motor 100 which operates under the control of unit 77. The motor is mounted on bar 96, as shown.

The structure just described, including worm screw 94 and motor 100, is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an y axis in the figure. As above, the structure functions in one mode to move the dispenser in y-axis increments having a selected distance in the range 5-250 μm, and in a second mode, to move the dispenser in precise y-axis increments of several microns (μm) or more, for positioning the dispenser at associated positions on adjacent supports.

The displacement assembly and structure for moving this assembly in the y axis are referred to herein collectively as positioning means for positioning the

dispensing device at a selected array position with respect to a support.

A holder 102 in the apparatus functions to hold a plurality of supports, such as supports 104 on which the microarrays of regent regions are to be formed by the apparatus. The holder provides a number of recessed slots, such as slot 106, which receive the supports, and position them at precise selected positions with respect to the frame bars on which the dispenser moving means is mounted.

As noted above, the control unit in the device functions to actuate the two stepper motors and dispenser solenoid in a sequence designed for automated operation of the apparatus in forming a selected microarray of reagent regions on each of a plurality of supports.

The control unit is constructed, according to conventional microprocessor control principles, to provide appropriate signals to each of the solenoid and each of the stepper motors, in a given timed sequence and for appropriate signalling time. The construction of the unit, and the settings that are selected by the user to achieve a desired array pattern, will be understood from the following description of a typical apparatus operation.

Initially, one or more supports are placed in one or more slots in the holder. The dispenser is then moved to a position directly above a well (not shown) containing a solution of the first reagent to be dispensed on the support(s). The dispenser solenoid is actuated now to lower the dispenser tip into this well, causing the capillary channel in the dispenser to fill. Motors 82, 100 are now actuated to position the dispenser at a selected array position at the first of the supports. Solenoid actuation of the dispenser is

then effective to dispense a selected-volume droplet of that reagent at this location. As noted above, this operation is effective to dispense a selected volume preferably between 2 pl and 2 nl of the reagent solution.

The dispenser is now moved to the corresponding position at an adjacent support and a similar volume of the solution is dispensed at this position. The process is repeated until the reagent has been dispensed at this preselected corresponding position on each of the supports.

Where it is desired to dispense a single reagent at more than two array positions on a support, the dispenser may be moved to different array positions at each support, before moving the dispenser to a new support, or solution can be dispensed at individual positions on each support, at one selected position, then the cycle repeated for each new array position.

To dispense the next reagent, the dispenser is positioned over a wash solution (not shown), and the dispenser tip is dipped in and out of this solution until the reagent solution has been substantially washed from the tip. Solution can be removed from the tip, after each dipping, by vacuum, compressed air spray, sponge, or the like.

The dispenser tip is now dipped in a second reagent well, and the filled tip is moved to a second selected array position in the first support. The process of dispensing reagent at each of the corresponding second-array positions is then carried as above. This process is repeated until an entire microarray of reagent solutions on each of the supports has been formed.

## IV. Microarray Substrate

This section describes embodiments of a substrate having a microarray of biological polymers carried on the substrate surface. Subsection A describes a multi-cell substrate, each cell of which contains a microarray, and preferably an identical microarray, of distinct biopolymers, such as distinct polynucleotides, formed on a porous surface. Subsection B describes a microarray of distinct polynucleotides bound on a glass slide coated with a polycationic polymer.

A. Multi-Cell Substrate

Fig. 9 illustrates, in plan view, a substrate 110 constructed according to the invention. The substrate has an 8 × 12 rectangular array 112 of cells, such as cells 114, 116, formed on the substrate surface. With reference to Fig. 10, each cell, such as cell 114, in turn supports a microarray 118 of distinct biopolymers, such as polypeptides or polynucleotides at known, addressable regions of the microarray. Two such regions forming the microarray are indicated at 120, and correspond to regions, such as regions 42, forming the microarray of distinct biopolymers shown in Fig. 3.

The 96-cell array shown in Fig. 9 has typically array dimensions between about 12 and 244 mm in width and 8 and 400 mm in length, with the cells in the array having width and length dimension of 1/12 and 1/8 the array width and length dimensions, respectively, *i.e.*, between about 1 and 20 in width and 1 and 50 mm in length.

The construction of substrate is shown cross-sectionally in Fig. 11, which is an enlarged sectional view taken along view line 124 in Fig. 9. The substrate includes a water-impermeable backing 126, such as a glass slide or rigid polymer sheet. Formed on the surface of the backing is a water-permeable film

128. The film is formed of a porous membrane material, such as nitrocellulose membrane, or a porous web material, such as a nylon, polypropylene, or PVDF porous polymer material. The thickness of the film is preferably between about 10 and 1000 μm. The film may be applied to the backing by spraying or coating uncured material on the backing, or by applying a preformed membrane to the backing. The backing and film may be obtained as a preformed unit from commercial source, *e.g.*, a plastic-backed nitrocellulose film available from Schleicher and Schuell Corporation.

With continued reference to Fig. 11, the film-covered surface in the substrate is partitioned into a desired array of cells by water-impermeable grid lines, such as lines 130, 132, which have infiltrated the film down to the level of the backing, and extend above the surface of the film as shown, typically a distance of 100 to 2000 μm above the film surface.

The grid lines are formed on the substrate by laying down an uncured or otherwise flowable resin or elastomer solution in an array grid, allowing the material to infiltrate the porous film down to the backing, then curing or otherwise hardening the grid lines to form the cell-array substrate.

One preferred material for the grid is a flowable silicone available from Loctite Corporation. The barrier material can be extruded through a narrow syringe (*e.g.*, 22 gauge) using air pressure or mechanical pressure. The syringe is moved relative to the solid support to print the barrier elements as a grid pattern. The extruded bead of silicone wicks into the pores of the solid support and cures to form a shallow waterproof barrier separating the regions of the solid support.

In alternative embodiments, the barrier element can be a wax-based material or a thermoset material such as epoxy. The barrier material can also be a UV-curing polymer which is exposed to UV light after being printed onto the solid support. The barrier material may also be applied to the solid support using printing techniques such as silk-screen printing. The barrier material may also be a heat-seal stamping of the porous solid support which seals its pores and forms a water-impervious barrier element. The barrier material may also be a shallow grid which is laminated or otherwise adhered to the solid support.

In addition to plastic-backed nitrocellulose, the solid support can be virtually any porous membrane with or without a non-porous backing. Such membranes are readily available from numerous vendors and are made from nylon, PVDF, polysulfone and the like. In an alternative embodiment, the barrier element may also be used to adhere the porous membrane to a non-porous backing in addition to functioning as a barrier to prevent cross contamination of the assay reagents.

In an alternative embodiment, the solid support can be of a non-porous material. The barrier can be printed either before or after the microarray of biomolecules is printed on the solid support.

As can be appreciated, the cells formed by the grid lines and the underlying backing are water-impermeable, having side barriers projecting above the porous film in the cells. Thus, defined-volume samples can be placed in each well without risk of cross-contamination with sample material in adjacent cells. In Fig. 11, defined volumes samples, such as sample 134, are shown in the cells.

As noted above, each well contains a microarray of distinct biopolymers. In one general embodiment, the

microarrays in the well are identical arrays of distinct biopolymers, *e.g.*, different sequence polynucleotides. Such arrays can be formed in accordance with the methods described in Section II, by depositing a first selected polynucleotide at the same selected microarray position in each of the cells, then depositing a second polynucleotide at a different microarray position in each well, and so on until a complete, identical microarray is formed in each cell.

In a preferred embodiment, each microarray contains about $10^3$ distinct polynucleotide or polypeptide biopolymers per surface area of less than about 1 $cm^2$. Also in a preferred embodiment, the biopolymers in each microarray region are present in a defined amount between about 0.1 femtomoles and 100 nanomoles. The ability to form high-density arrays of biopolymers, where each region is formed of a well-defined amount of deposited material, can be achieved in accordance with the microarray-forming method described in Section II.

Also in a preferred embodiments, the biopolymers are polynucleotides having lengths of at least about 50 bp, *i.e.*, substantially longer than oligonucleotides which can be formed in high-density arrays by schemes involving parallel, step-wise polymer synthesis on the array surface.

In the case of a polynucleotide array, in an assay procedure, a small volume of the labeled DNA probe mixture in a standard hybridization solution is loaded onto each cell. The solution will spread to cover the entire microarray and stop at the barrier elements. The solid support is then incubated in a humid chamber at the appropriate temperature as required by the assay.

Each assay may be conducted in an "open-face" format where no further sealing step is required, since the hybridization solution will be kept properly hydrated by the water vapor in the humid chamber. At the conclusion of the incubation step, the entire solid support containing the numerous microarrays is rinsed quickly enough to dilute the assay reagents so that no significant cross contamination occurs. The entire solid support is then reacted with detection reagents if needed and analyzed using standard colorimetric, radioactive or fluorescent detection means. All processing and detection steps are performed simultaneously to all of the microarrays on the solid support ensuring uniform assay conditions for all of the microarrays on the solid support.

B. Glass-Slide Polynucleotide Array

Fig. 5 shows a substrate 136 formed according to another aspect of the invention, and intended for use in detecting binding of labeled polynucleotides to one or more of a plurality distinct polynucleotides. The substrate includes a glass substrate 138 having formed on its surface, a coating of a polycationic polymer, preferably a cationic polypeptide, such as polylysine or polyarginine. Formed on the polycationic coating is a microarray 140 of distinct polynucleotides, each localized at known selected array regions, such as regions 142.

The slide is coated by placing a uniform-thickness film of a polycationic polymer, *e.g.*, poly-l-lysine, on the surface of a slide and drying the film to form a dried coating. The amount of polycationic polymer added is sufficient to form at least a monolayer of polymers on the glass surface. The polymer film is bound to surface via electrostatic binding between

negative silyl-OH groups on the surface and charged amine groups in the polymers. Poly-l-lysine coated glass slides may be obtained commercially, e.*g*., from Sigma Chemical Co. (St. Louis, MO).

To form the microarray, defined volumes of distinct polynucleotides are deposited on the polymer-coated slide, as described in Section II. According to an important feature of the substrate, the deposited polynucleotides remain bound to the coated slide surface non-covalently when an aqueous DNA sample is applied to the substrate under conditions which allow hybridization of reporter-labeled polynucleotides in the sample to complementary-sequence (single-stranded) polynucleotides in the substrate array. The method is illustrated in Examples 1 and 2.

To illustrate this feature, a substrate of the type just described, but having an array of same-sequence polynucleotides, was mixed with fluorescent-labeled complementary DNA under hybridization conditions. After washing to remove non-hybridized material, the substrate was examined by low-power fluorescence microscopy. The array can be visualized by the relatively uniform labeling pattern of the array regions.

In a preferred embodiment, each microarray contains at least $10^3$ distinct polynucleotide or polypeptide biopolymers per surface area of less than about 1 $cm^2$. In the embodiment shown in Fig. 5, the microarray contains 400 regions in an area of about 16 $mm^2$, or $2.5 \times 10^3$ regions/$cm^2$. Also in a preferred embodiment, the polynucleotides in the each microarray region are present in a defined amount between about 0.1 femtomoles and 100 nanomoles in the case of polynucleotides. As above, the ability to form high-

density arrays of this type, where each region is formed of a well-defined amount of deposited material, can be achieved in accordance with the microarray-forming method described in Section II.

Also in a preferred embodiments, the polynucleotides have lengths of at least about 50 bp, *i.e.*, substantially longer than oligonucleotides which can be formed in high-density arrays by various in situ synthesis schemes.

V. Utility

Microarrays of immobilized nucleic acid sequences prepared in accordance with the invention can be used for large scale hybridization assays in numerous genetic applications, including genetic and physical mapping of genomes, monitoring of gene expression, DNA sequencing, genetic diagnosis, genotyping of organisms, and distribution of DNA reagents to researchers.

For gene mapping, a gene or a cloned DNA fragment is hybridized to an ordered array of DNA fragments, and the identity of the DNA elements applied to the array is unambiguously established by the pixel or pattern of pixels of the array that are detected. One application of such arrays for creating a genetic map is described by Nelson, *et al.* (1993). In constructing physical maps of the genome, arrays of immobilized cloned DNA fragments are hybridized with other cloned DNA fragments to establish whether the cloned fragments in the probe mixture overlap and are therefore contiguous to the immobilized clones on the array. For example, Lehrach, *et al.*, describe such a process.

The arrays of immobilized DNA fragments may also be used for genetic diagnostics. To illustrate, an array containing multiple forms of a mutated gene or genes can be probed with a labeled mixture of a

patient's DNA which will preferentially interact with only one of the immobilized versions of the gene.

The detection of this interaction can lead to a medical diagnosis. Arrays of immobilized DNA fragments can also be used in DNA probe diagnostics. For example, the identity of a pathogenic microorganism can be established unambiguously by hybridizing a sample of the unknown pathogen's DNA to an array containing many types of known pathogenic DNA. A similar technique can also be used for unambiguous genotyping of any organism. Other molecules of genetic interest, such as cDNA's and RNA's can be immobilized on the array or alternately used as the labeled probe mixture that is applied to the array.

In one application, an array of cDNA clones representing genes is hybridized with total cDNA from an organism to monitor gene expression for research or diagnostic purposes. Labeling total cDNA from a normal cell with one color fluorophore and total cDNA from a diseased cell with another color fluorophore and simultaneously hybridizing the two cDNA samples to the same array of cDNA clones allows for differential gene expression to be measured as the ratio of the two fluorophore intensities. This two-color experiment can be used to monitor gene expression in different tissue types, disease states, response to drugs, or response to environmental factors. & An example of this approach is illustrated in Examples 2, described with respect to Fig. 8.

By way of example and without implying a limitation of scope, such a procedure could be used to simultaneously screen many patients against all known mutations in a disease gene. This invention could be used in the form of, for example, 96 identical 0.9 cm × 2.2 cm microarrays fabricated on a single 12 cm × 18 cm

sheet of plastic-backed nitrocellulose where each microarray could contain, for example, 100 DNA fragments representing all known mutations of a given gene. The region of interest from each of the DNA samples from 96 patients could be amplified, labeled, and hybridized to the 96 individual arrays with each assay performed in 100 microliters of hybridization solution. The approximately 1 thick silicone rubber barrier elements between individual arrays prevent cross contamination of the patient samples by sealing the pores of the nitrocellulose and by acting as a physical barrier between each microarray. The solid support containing all 96 microarrays assayed with the 96 patient samples is incubated, rinsed, detected and analyzed as a single sheet of material using standard radioactive, fluorescent, or colorimetric detection means (Maniatas, *et al.*, 1989). Previously, such a procedure would involve the handling, processing and tracking of 96 separate membranes in 96 separate sealed chambers. By processing all 96 arrays as a single sheet of material, significant time and cost savings are possible.

The assay format can be reversed where the patient or organism's DNA is immobilized as the array elements and each array is hybridized with a different mutated allele or genetic marker. The gridded solid support can also be used for parallel non-DNA ELISA assays. Furthermore, the invention allows for the use of all standard detection methods without the need to remove the shallow barrier elements to carry out the detection step.

In addition to the genetic applications listed above, arrays of whole cells, peptides, enzymes, antibodies, antigens, receptors, ligands, phospholipids, polymers, drug cogener preparations or

chemical substances can be fabricated by the means described in this invention for large scale screening assays in medical diagnostics, drug discovery, molecular biology, immunology and toxicology.

The multi-cell substrate aspect of the invention allows for the rapid and convenient screening of many DNA probes against many ordered arrays of DNA fragments. This eliminates the need to handle and detect many individual arrays for performing mass screenings for genetic research and diagnostic applications. Numerous microarrays can be fabricated on the same solid support and each microarray reacted with a different DNA probe while the solid support is processed as a single sheet of material.

The following examples illustrate, but in no way are intended to limit, the present invention.

## Example 1

### Genomic-Complexity Hybridization to Micro DNA Arrays Representing the Yeast *Saccharomyces cerevisiae* Genome with Two-Color Fluorescent Detection

The array elements were randomly amplified PCR (Bohlander, *et al.*, 1992) products using physically mapped lambda clones of *S. cerevisiae* genomic DNA templates (Riles, *et al.*, 1993). The PCR was performed directly on the lambda phage lysates resulting in an amplification of both the 35 kb lambda vector and the 5-15 kb yeast insert sequences in the form of a uniform distribution of PCR product between 250-1500 base pairs in length. The PCR product was purified using Sephadex G50 gel filtration (Pharmacia, Piscataway, NJ) and concentrated by evaporation to dryness at room temperature overnight. Each of the 864 amplified

lambda clones was rehydrated in 15 $\mu$l of 3 × SSC in preparation for spotting onto the glass.

The micro arrays were fabricated on microscope slides which were coated with a layer of poly-l-lysine (Sigma). The automated apparatus described in Section IV loaded 1 $\mu$l of the concentrated lambda clone PCR product in 3 × SSC directly from 96 well storage plates into the open capillary printing element and deposited ~5 nl of sample per slide at 380 micron spacing between spots, on each of 40 slides. The process was repeated for all 864 samples and 8 control spots. After the spotting operation was complete, the slides were rehydrated in a humid chamber for 2 hours, baked in a dry 80° vacuum oven for 2 hours, rinsed to remove un-absorbed DNA and then treated with succinic anhydride to reduce non-specific adsorption of the labeled hybridization probe to the poly-l-lysine coated glass surface. Immediately prior to use, the immobilized DNA on the array was denatured in distilled water at 90° for 2 minutes.

For the pooled chromosome experiment, the 16 chromosomes of *Saccharomyces cerevisiae* were separated in a CHEF agarose gel apparatus (Biorad, Richmond, CA). The six largest chromosomes were isolated in one gel slice and the smallest 10 chromosomes in a second gel slice. The DNA was recovered using a gel extraction kit (Qiagen, Chatsworth, CA). The two chromosome pools were randomly amplified in a manner similar to that used for the target lambda clones. Following amplification, 5 micrograms of each of the amplified chromosome pools were separately random-primer labeled using Klenow polymerase (Amersham, Arlington Heights, IL) with a lissamine conjugated nucleotide analog (Dupont NEN, Boston, MA) for the pool containing the six largest chromosomes, and with a fluorescein

conjugated nucleotide analog (BMB) for the pool containing smallest ten chromosomes. The two pools were mixed and concentrated using an ultrafiltration device (Amicon, Danvers, MA).

Five micrograms of the hybridization probe consisting of both chromosome pools in 7.5 $\mu$l of TE was denatured in a boiling water bath and then snap cooled on ice. 2.5 $\mu$l of concentrated hybridization solution (5 × SSC and 0.1% SDS) was added and all 10 $\mu$l transferred to the array surface, covered with a cover slip, placed in a custom-built single-slide humidity chamber and incubated at 60° for 12 hours. The slides were then rinsed at room temperature in 0.1 × SSC and 0.1%SDS for 5 minutes, cover slipped and scanned.

A custom built laser fluorescent scanner was used to detect the two-color hybridization signals from the 1.8 × 1.8 cm array at 20 micron resolution. The scanned image was gridded and analyzed using custom image analysis software. After correcting for optical crosstalk between the fluorophores due to their overlapping emission spectra, the red and green hybridization values for each clone on the array were correlated to the known physical map position of the clone resulting in a computer-generated color karyotype of the yeast genome.

Figure 6 shows the hybridization pattern of the two chromosome pools. A red signal indicates that the lambda clone on the array surface contains a cloned genomic DNA segment from one of the largest six yeast chromosomes. A green signal indicates that the lambda clone insert comes from one of the smallest ten yeast chromosomes. Orange signals indicate repetitive sequences which cross hybridized to both chromosome pools. Control spots on the array confirm that the hybridization is specific and reproducible.

The physical map locations of the genomic DNA fragments contained in each of the clones used as array elements have been previously determined by Olson and co-workers (Riles, *et al.*) allowing for the automatic generation of the color karyotype shown in Figure 7. The color of a chromosomal section on the karyotype corresponds to the color of the array element containing the clone from that section. The black regions of the karyotype represent false negative dark spots on the array (10%) or regions of the genome not covered by the Olson clone library (90%). Note that the largest six chromosomes are mainly red while the smallest ten chromosomes are mainly green matching the original CHEF gel isolation of the hybridization probe. Areas of the red chromosomes containing green spots and vice-versa are probably due to spurious sample tracking errors in the formation of the original library and in the amplification and spotting procedures.

The yeast genome arrays have also been probed with individual clones or pools of clones that are fluorescently labeled for physical mapping purposes. The hybridization signals of these clones to the array were translated into a position on the physical map of yeast.

## Example 2

## Total cDNA Hybridized to Micro Arrays of cDNA Clones with Two-Color Fluorescent Detection

24 clones containing cDNA inserts from the plant *Arabidopsis* were amplified using PCR. Salt was added to the purified PCR products to a final concentration of 3 × SSC. The cDNA clones were spotted on poly-l-lysine coated microscope slides in a manner similar to Example 1. Among the cDNA clones was a clone

representing a transcription factor HAT 4, which had previously been used to create a transgenic line of the plant *Arabidopsis*, in which this gene is present at ten times the level found in wild-type *Arabidopsis* (Schena, *et al.*, 1992).

Total poly-A mRNA from wild type *Arabidopsis* was isolated using standard methods (Maniatis, *et al.*, 1989) and reverse transcribed into total cDNA, using fluorescein nucleotide analog to label the cDNA product (green fluorescence). A similar procedure was performed with the transgenic line of *Arabidopsis* where the transcription factor HAT4 was inserted into the genome using standard gene transfer protocols. cDNA copies of mRNA from the transgenic plant are labeled with a lissamine nucleotide analog (red fluorescence). Two micrograms of the cDNA products from each type of plant were pooled together and hybridized to the cDNA clone array in a 10 microliter hybridization reaction in a manner similar to Example 1. Rinsing and detection of hybridization was also performed in a manner similar to Example 1. Fig. 8 show the resulting hybridization pattern of the array.

Genes equally expressed in wild type and the transgenic *Arabidopsis* appeared yellow due to equal contributions of the green and red fluorescence to the final signal. The dots are different intensities of yellow indicating various levels of gene expression. The cDNA clone representing the transcription factor HAT4, expressed in the transgenic line of *Arabidopsis* but not detectably expressed in wild type *Arabidopsis*, appears as a red dot (with the arrow pointing to it), indicating the preferential expression of the transcription factor in the red-labeled transgenic *Arabidopsis* and the relative lack of expression of the

transcription factor in the green-labeled wild type *Arabidopsis*.

An advantage of the microarray hybridization format for gene expression studies is the high partial concentration of each cDNA species achievable in the 10 microliter hybridization reaction. This high partial concentration allows for detection of rare transcripts without the need for PCR amplification of the hybridization probe which may bias the true genetic representation of each discrete cDNA species.

Gene expression studies such as these can be used for genomics research to discover which genes are expressed in which cell types, disease states, development states or environmental conditions. Gene expression studies can also be used for diagnosis of disease by empirically correlating gene expression patterns to disease states.

## Example 3

## Multiplexed Colorimetric Hybridization on a Gridded Solid Support

A sheet of plastic-backed nitrocellulose was gridded with barrier elements made from silicone rubber according to the description in Section IV-A. The sheet was soaked in 10 × SSC and allowed to dry. As shown in Fig. 12, 192 M13 clones each with a different yeast inserts were arrayed 400 microns apart in four quadrants of the solid support using the automated device described in Section III. The bottom left quadrant served as a negative control for hybridization while each of the other three quadrants was hybridized simultaneously with a different oligonucleotide using the open-face hybridization technology described in Section IV-A. The first two and last four elements of

each array are positive controls for the colorimetric detection step.

The oligonucleotides were labeled with fluorescein which was detected using an anti-fluorescein antibody conjugated to alkaline phosphatase that precipitated an NBT/BCIP dye on the solid support (Amersham). Perfect matches between the labeled oligos and the M13 clones resulted in dark spots visible to the naked eye and detected using an optical scanner (HP ScanJet II) attached to a personal computer. The hybridization patterns are different in every quadrant indicating that each oligo found several unique M13 clones from among the 192 with a perfect sequence match. Note that the open capillary printing tip leaves detectable dimples on the nitrocellulose which can be used to automatically align and analyze the images.

Although the invention has been described with respect to specific embodiments and methods, it will be clear that various changes and modification may be made without departing from the invention.

# Specification as originally filed and downloaded from PAIR Serial No. 08/514,875

Attorney Docket No.: 8600-0138.30 Patent

**Transmittal of Utility Patent Application for Filing**

*Certification Under 37 C.F.R. §1.10 (if applicable)*

| TB 585 979 097 US | 14 August 1995 |
|---|---|
| "Express Mail" Label Number | Date of Deposit |

I hereby certify that this Transmittal letter, enclosed application, and any other documents referred to as enclosed herein are being deposited in an envelope with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231

Dari Shalon — (Print Name of Person Mailing Application)

(Signature of Person Mailing Application)

# METHOD AND GENE-ARRAY DEVICE FOR ANALYZING GENE EXPRESSION PATTERNS

This invention is a continuation-in part of U.S. patent application Serial No. 08/477,809 for Method and Apparatus for Fabricating Microarrays of Biological Samples, filed June 7, 1995, and now allowed, which is a continuation-in-part of U.S. patent application Serial No. 08/261,388 for Method and Apparatus for Fabricating Microarrays of Biological Samples, filed June 17, 1994. These two applications are incorporated herein by reference.

## Field of the Invention

This invention relates to a method and gene-array device for detecting and monitoring gene expression levels specifically related to a given disease-related state, and to a method for constructing the gene-array device.

## References

Ausubel, F.M., *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley and Sons, Inc., Media, PA).

Cole, C.G., *et al.*, *Genomics* 14:931-8 (1992).

Diamandis E.P., *Critical Rev in Clinical Laboratory Sci, 1992*, 29(3-4):269-305.

Lehrach, *et al.*, HYBRIDIZATION FINGERPRINTING IN GENOME MAPPING AND SEQUENCING, GENOME ANALYSIS, VOL 1 (Davies and Tilgham, Eds.), Cold Spring Harbor Press, pp. 39-81 (1990).

Lytras A., *et al.*, *Endocrinology*, 1994 Jun, 134(6):2461-7.

Maniatis, *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press (1989).

Manoni, M., *et al.*, *Biotechniques* 12:48-50, 52-3 (1992).

Maser R.L., *et al.*, *Seminars in Nephrology*, 1995 Jan, 15(1):29-42.

Mullis, K.B., U.S. Patent No. 4,683,202, issued 28 July 1987a.

Mullis, K.B., *et al.*, U.S. Patent No. 4,683,195, issued 28 July 1987b.

Nelson, *et al.*, *Nature Genetics* 4:11-18 (1993).

Orr S.L., *et al.*, *Proc Nat Acad Sci, USA*, 1994 Dec 6, 91(25):11869-73.

Osterland C.K., *Clinical Chemistry*, 1994 Nov, 40(11 Pt 2):2146-53.

Piatak M. Jr, *et al.*, *Science*, 1993 Mar 19, 259(5102):1749-54.

Sambrook J., *et al.*, *Molecular Cloning, A Laboratory Manual*, Second Edition, Cold Spring Harbor Press, 1989

Schena, M. *et al.*, *Proc. Nat. Acad. Sci. USA* 89:3894-3898 (1992).

## Background of the Invention

A variety of methods for analyzing gene products and gene expression are available. Northern blot is

one widely used method for monitoring gene expression (Sambrook et al). In this method, a cellular RNA fraction, typically total mRNA, is electrophoretically separated on a gel and the separated RNA species are transferred to a blot substrate. The RNA species are then hybridized with one or more labeled probes of interest, e.g., a labeled RNA or cDNA fragment on the substrate. The purpose of the method is to detect the size range and/or relative quantity of an expressed RNA species complementary to the labeled probe.

In general, sensitivity in the Northern blot is limited to mRNA species present at a level of about 1:10,000 of the total mRNA using radioactive detection, and Northern blots require up to 50 micrograms of mRNA per lane. Northern blots are labor intensive and in general not well suited to diagnostic applications.

The amplification step often required in Northern analysis may bias the relative abundance of discrete cDNA species in a complex mixture. Although methods which allow for reliable quantitation of the RT-PCR process (Piatak, et al) have been proposed, these are not practical for analysis of large numbers of genes.

It has also been proposed to analyze patterns of gene expression by hybridizing a gene of interest to colony blots of different cDNA libraries, where the frequency of hits provides a measure of "differential display" of the gene in the different tissues from which the cDNA libraries were originally made (Maser et al). This method requires the fabrication of a cDNA library and the labor-intensive analysis of unordered dot blots from the colony hybridizations.

Recently, there has been considerable interest in monitoring gene expression using tag sequencing. Here a cDNA library from a specific tissue type or disease state is made from poly A mRNA. Individual cDNA clones

with inserts of 1-2 kb are selected at random and a "tag" of around 200 bases of each cDNA insert is sequenced. An expression profile is generated in the form of a computer database of the tag sequences for thousands of cDNA clones from the cDNA library (Orr). Computer analysis of expression profiles can determine which genes are differentially expressed in a specific tissue type or disease state. To date, however, the use of tag sequences has been limited by lack of information about the functional roles of most tag sequences, and in fact, assigning functional roles to tag sequences represents one of the challenging problems of the Human Genome Project.

In view of these limitations in identifying and quantitating gene expression levels for large numbers of expressed genes, it has not been practical heretofore to employ multi-gene expression as a sensitive test for cell status, e.g., in a disease state, or as a method for monitoring the effect of therapeutic treatment on diseased-state cells.

## Summary of the Invention

The invention includes, in one aspect, a method of constructing a subarray of genes whose gene expression levels are specifically related to the differences between test cells relative to control cells. The method includes first obtaining and preparing reporter-labeled copies of messenger nucleic acid from control cells in a population of control individuals, and from test cells in a population of test individuals having a shared phenotype of interest, e.g., a disease state, that is not present in control individuals.

The reporter-labeled nucleic acid from test and control cells is applied to a substrate having an array

(e.g., a microarray) of at least $10^2$, and preferably $10^3$, distinct gene sequences. The nucleic acid is applied under conditions effective to hybridize the nucleic acid to complementary-sequence genes on the array.

The pattern of reporter levels for nucleic acids from the test cells is compared with that of nucleic acids from the control cells, and from this, the genes on the microarray which show a significant elevation or reduction in reporter levels, when compared with control levels, are identified. The array is formed with the identified genes.

The array is preferably a microarray formed on a single, contiguous substrate, at a density of at least $10^3$ distinct gene sequences per $cm^2$ surface area, and may contain $10^3$ to $10^4$ or more distinct gene sequences. Preferably, each distinct gene sequence is disposed at a separate, defined position in said array and is present in a defined amount between about 0.1 femtomole and 100 nanomoles.

The gene sequences forming the array may be obtained from a single tissue source or, preferably, from multiple tissue sources, and typically include cDNA sequences.

In one general embodiment, the test-cell nucleic acids from each test individual are applied to a separate array, and the genes of interest are identified by identifying those genes on the separate arrays which show a statistically significant elevation or reduction in reporter levels, when compared with control levels. In another general embodiment, the test-cell nucleic acids from the test individuals are pooled and applied to a single array.

The test- and control-cell nucleic acids may have different fluorescent reporters, allowing the nucleic acids from the test- and control cells to be applied to the same array. Alternatively, the test- and control-cell nucleic acids may be applied to different microarrays, with the same or different reporters.

In another aspect, the invention includes a method of detecting or monitoring the treatment status of a selected disease condition. In practicing the method, there is first prepared reporter-labeled copies of messenger nucleic acid obtained from test cells associated with the disease condition. The nucleic acid is applied to a subarray of genes which are characterized by a statistically significant increase or decrease in gene level expression, when compared with the level of gene expression in the same cell type or types in a control, non-disease state, under conditions effective to hybridize said nucleic acid species to complementary-sequence genes in said array. The levels of reporter associated with the genes in the array are determined, and from this, there is formed a pattern of gene expression. A comparison of the gene-expression pattern with a known pattern of gene expression associated with the disease condition, permits detection or monitoring of the treatment status of the disease state.

The method employs a gene-array device constructed according to another aspect of the invention. The device includes a substrate, and a subarray of genes which each show a statistically significant increase or a statistically significant decrease in gene ~~level~~ expression level when compared with the level of gene expression in a control cell type.

These and other objects and features of the invention will become more fully apparent when the

following detailed description of the invention is read in conjunction with the accompanying figures.

## Brief Description of the Drawings

Fig. 1 is a side view of a solution-dispensing device having a open-capillary dispensing head constructed for use in one embodiment of the invention;

Figs. 2A-2C illustrate steps in the delivery of a fixed-volume bead on a hydrophobic surface employing the dispensing head from Fig. 1, in accordance with one embodiment of the method of the invention;

Fig. 3 shows a portion of a two-dimensional array of analyte-assay regions constructed according to the method of the invention;

Fig. 4 is a planar view showing components of an automated apparatus for forming arrays in accordance with the invention.

Fig. 5 shows a fluorescent image of an actual 20 × 20 array of 400 fluorescently-labeled DNA samples immobilized on a poly-l-lysine coated slide, where the total area covered by the 400 element array is 16 square millimeters;

Fig. 6 is a fluorescent image of a 1.8 cm × 1.8 cm microarray containing lambda clones with yeast inserts, the fluorescent signal arising from the hybridization to the array with approximately half the yeast genome labeled with a green fluorophore and the other half with a red fluorophore;

Figs. 7A and 7B show scans of hybridization signals from an array of genes probed with fluorescently-labeled *Arabidopsis* cDNA at high (7A) and intermediate (7B) photomultiplier tube settings.

Figs. 8A and 8B show scans of hybridization signals from an array of genes probed with fluorescently-labeled *Arabidopsis* wild-type (8A) or

transgenic HAT4 (8B) cDNA at low photomultiplier tube settings.

Figs. 9A and 9B show scans of hybridization signals from an array of genes probed with fluorescently-labeled *Arabidopsis* wild-type root (9A) or wild-type leaf (9B) cDNA at intermediate photomultiplier tube settings.

## Detailed Description of the Invention

I. Definitions

Unless indicated otherwise, the terms defined below have the following meanings:

"Distinct gene sequences" or "different gene sequences", as applied to the gene sequences forming an array or microarray or subarray, refers to polynucleotides containing distinct, i.e., different, gene sequences. The different-sequence polynucleotides may be partially or completely sequenced, as with expressed-sequence tags (EST's) or unsequenced, as with an unsequenced cDNA library.

An "array of distinct gene sequences" refers to a linear or two-dimensional array of distinct gene sequences, where the array may also contain regions with different graded concentrations of same-sequence polynucleotides, and/or mixtures of two or more distinct-sequence polynucleotides.

A "microarray of distinct gene sequences" refers to an array having a density of distinct gene sequences of at least about 100/cm$^2$, for example about 400/cm$^2$, and preferably at least about 1000/cm$^2$. The regions in a microarray have typical dimensions, e.g., diameters, in the range of between about 10-500 μm, for example about 250 μm, and are separated from other regions in the array by about the same distance, and contain

typically, 0.1 femtomole to 100 ~~nanomole nucleic acid~~ nanomoles of nucleic acid molecules.

A "subarray of distinct gene sequences" is an array formed from a subset of gene sequences in a larger array. The subset is typically composed of gene sequences whose level of expression in cells in a selected physiological state or disease state obtained from a population of test individuals is significantly higher or lower than in control cells obtained from a population of control individuals.

"Cells of a given cell type or types" refers to cells obtained from one or more particular tissues or organs, e.g., hepatocytes, heart muscle cells, pancreatic cells, or non-differentiated embryonic tissue, or to a particular blood cell type or types, e.g., peripheral blood lymphocytes.

Cells having a "selected physiological state or disease condition" or "test cells" refer to cells of a given cell type or types which are (i) in a defined state of differentiation or activation, e.g., by gene activation, (ii) infected by a defined infectious agent, e.g., HIV-infected T cells, (iii) in a neoplastic state, i.e., tumor cells, (iv) in a chemical- or physical-response state, i.e., after exposure to a pharmacological agent with respect to control cells of the same type or types.

Cells of the same cell type or types as test cells, but which are (i) in a non-differentiated or non-activated state, (ii) uninfected, (iii) in a normal, non-neoplastic state, or (iv) in a control, no-drug state, are referred to herein as control cells.

A "population of test individuals" includes at least 5, and preferably 50 or more individuals all of whom share a common phenotype related to the individuals' test cells, e.g., individuals who all have

a common disease or are infected by the same infectious agent.

A "population of control individuals" includes at least 5, and preferably 50 or more individuals all of whom share a common control phenotype related to the individuals' control cells, e.g., normal, disease-free and/or drug-free individuals, or share the same physiological state of other phenotype of interest.

"Reporter-labeled copies of messenger nucleic acid" refers to reporter-labeled mRNA transcripts obtained from test or control cells or cDNAs produced from such transcripts. The reporter label is any detectable reporter, and typically a fluorescent reporter.

A support surface of an array is "hydrophobic" if a aqueous-medium droplet applied to the surface does not spread out substantially beyond the area size of the applied droplet. That is, the surface acts to prevent spreading of the droplet applied to the surface by hydrophobic interaction with the droplet.

A "meniscus" means a concave or convex surface that forms on the bottom of a liquid in a channel as a result of the surface tension of the liquid.

## II. Method of Microarray Formation

This section describes a method of forming a microarray of distinct gene sequences on a solid support or substrate, for use in the method of the invention.

Fig. 1 illustrates, in a partially schematic view, a solution-dispensing device 10 useful in producing such a microarray. The device generally includes a dispenser 12 having an elongate open capillary channel 14 adapted to hold a quantity of the solution of a given-sequence gene or gene region, such as indicated

at 16. The capillary channel is formed by a pair of spaced-apart, coextensive, elongate members 12a, 12b which are tapered toward one another and converge at a tip region 18 at the lower end of the channel.

With continued reference to Fig. 1, the dispenser device also includes structure for moving the dispenser rapidly toward and away from a support surface, for effecting deposition of a known amount of solution in the dispenser on a support, as will be described below with reference to Figs. 2A-2C. In the embodiment shown, this structure includes a solenoid 22 which is activatable to draw a solenoid piston 24 rapidly downwardly, then release the piston, e.g., under spring bias, to a normal, raised position, as shown. The dispenser is carried on the piston by a connecting member 26, as shown.

The dispensing device just described is carried on an arm 28 that may be moved either linearly or in an x-y plane to position the dispenser at a selected deposition position, as will be described.

Figs. 2A-2C illustrate the method of depositing a known amount of solution of a distinct-sequence polynucleotide solution in the just-described dispenser on the surface of a solid support, such as the support indicated at 30. The support is a polymer, glass, or other solid-material support having a surface indicated at 31.

In one general embodiment, the surface is a relatively hydrophilic, *i.e.*, wettable surface, such as a surface having native, bound or covalently attached charged groups. One such surface described below is a glass surface having an absorbed layer of a polycationic polymer, such as poly-l-lysine.

Initially, the dispenser is loaded with molecules of a selected gene sequence, such as by dipping the

dispenser tip, after washing, into a solution of the gene sequence solution, and allowing filling by capillary flow into the dispenser channel. The dispenser is now moved to a selected position with respect to a support surface, placing the dispenser tip directly above the support-surface position at which the polynucleotide solution is to be deposited. This movement takes place with the dispenser tip in its raised position, as seen in Fig. 2A, where the tip is typically at least several (1-5) mm above the surface of the substrate.

With the dispenser so positioned, solenoid 22 is now activated to cause the dispenser tip to move rapidly toward and away from the substrate surface, making momentary contact with the surface, in effect, tapping the tip of the dispenser against the support surface. The tapping movement of the tip against the surface acts to break the liquid meniscus in the tip channel, bringing the liquid in the tip into contact with the support surface. This, in turn, produces a flowing of the liquid into the capillary space between the tip and the surface, acting to draw liquid out of the dispenser channel, as seen in Fig. 2B.

Fig. 2C shows flow of fluid from the tip onto the support surface, which in this case is a hydrophobic surface. The figure illustrates that liquid continues to flow from the dispenser onto the support surface until it forms a liquid bead 32. At a given bead size, *i.e.*, volume, the tendency of liquid to flow onto the surface will be balanced by the hydrophobic surface interaction of the bead with the support surface, which acts to limit the total bead area on the surface, and by the surface tension of the droplet, which tends toward a given bead curvature. At this point, a given bead volume will have formed, and continued contact of

the dispenser tip with the bead, as the dispenser tip is being withdrawn, will have little or no effect on bead volume.

For liquid-dispensing on a more hydrophilic surface, the liquid will have less of a tendency to bead, and the dispensed volume will be more sensitive to the total dwell time of the dispenser tip in the immediate vicinity of the support surface, *e.g.*, the positions illustrated in Figs. 2B and 2C.

The desired deposition volume, *i.e.*, bead volume, formed by this method is preferably in the range 2 pl (picoliters) to 2 nl (nanoliters), although volumes as high as 100 nl or more may be dispensed. It will be appreciated that the selected dispensed volume will depend on (i) the "footprint" of the dispenser tip, *i.e.*, the size of the area spanned by the tip, (ii) the hydrophobicity of the support surface, and (iii) the time of contact with and rate of withdrawal of the tip from the support surface. In addition, bead size may be reduced by increasing the viscosity of the medium, effectively reducing the flow time of liquid from the dispenser onto the support surface. The drop size may be further constrained by depositing the drop in a hydrophilic region surrounded by a hydrophobic grid pattern on the support surface.

In a typical embodiment, the dispenser tip is tapped rapidly against the support surface, with a total residence time in contact with the support of less than about 1 msec, and a rate of upward travel from the surface of about 10 cm/sec.

Assuming that the bead that forms on contact with the surface is a hemispherical bead, with a diameter approximately equal to the width of the dispenser tip, as shown in Fig. 2C, the volume of the bead formed in relation to dispenser tip width (d) is given in Table 1

below. As seen, the volume of the bead ranges between 2 pl to 2 nl as the width size is increased from about 20 to 200 μm.

**Table 1**

| d | Volume (nl) |
|---|---|
| 20 μm | $2 \times 10^{-3}$ |
| 50 μm | $3.1 \times 10^{-2}$ |
| 100 μm | $2.5 \times 10^{-1}$ |
| 200 μm | 2 |

At a given tip size, bead volume can be reduced in a controlled fashion by increasing surface hydrophobicity, reducing time of contact of the tip with the surface, increasing rate of movement of the tip away from the surface, and/or increasing the viscosity of the medium. Once these parameters are fixed, a selected deposition volume in the desired pl to nl range can be achieved in a repeatable fashion.

After depositing a bead at one selected location on a support, the tip is typically moved to a corresponding position on a second support, a droplet is deposited at that position, and this process is repeated until a liquid droplet of the polynucleotide solution has been deposited at a selected position on each of a plurality of supports.

The tip is then washed to remove the polynucleotide solution, filled with another distinct-sequence polynucleotide solution and this solution is now deposited at another array position on each of the supports. In one embodiment, the tip is washed and refilled by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and

(iii) dipping the capillary channel into the new gene-sequence solution.

The tweezers-like, open-capillary dispenser tip provides the advantages that (i) the open channel of the tip facilitates rapid, efficient washing and drying before reloading the tip with a new, e.g., different-sequence polynucleotide solution, (ii) passive capillary action can load the sample directly from a standard microwell plate while retaining sufficient sample in the open capillary reservoir for the printing of numerous arrays, (iii) open capillaries are less prone to clogging than closed capillaries, and (iv) open capillaries do not require a perfectly faced bottom surface for fluid delivery.

A portion of a microarray 36 formed on the surface 38 of a solid support 40 in accordance with the method just described is shown in Fig. 3. The array is formed of a plurality of distinct-sequence gene regions, such as regions 42, where each region may include a different gene sequence or different concentration of gene sequence. As indicated above, the diameter of each region is preferably between about 20-200 μm, where each region contains between about 0.1 femtomole to 100 nanomoles of the distinct-sequence polynucleotide.

The spacing between each region and its closest (non-diagonal) neighbor, measured from center-to-center (indicated at 44), is preferably in the range of about 20-400 μm. Thus, for example, an array having a center-to-center spacing of about 250 μm contains about 40 regions/cm or 1,600 regions/$cm^2$. After formation of the array, the support is treated to evaporate the liquid of the droplet forming each region, to leave a

desired array of dried, relatively flat regions. This drying may be done by heating or under vacuum.

In some cases, it is desired to first rehydrate the droplets containing the polynucleotides to allow for more time for adsorption to the solid support. It is also possible to spot out the polynucleotides in a humid environment so that droplets do not dry until the arraying operation is complete, or to spot out the polynucleotides on a heated surface to increase the rate of absorption.

Fig. 4 shows, in simplified form, portions of an apparatus designed to automated production of a large number of microarrays of the types described above. A dispenser device 72 in the apparatus has the basic construction described above with respect to Fig. 1, and includes a dispenser 74 having an open-capillary channel terminating at a tip, substantially as shown in Figs. 1 and 2A-2C.

The dispenser is mounted in the device for movement toward and away from a dispensing position at which the tip of the dispenser taps a support surface, to dispense a selected volume of polynucleotide solution, as described above. This movement is effected by a solenoid 76 as described above. Solenoid 76 is under the control of a control unit 77 whose operation will be described below.

The dispenser device is carried on an arm 74 which is threadedly mounted on a worm screw 80 driven (rotated) in a desired direction by a stepper motor 82 also under the control of unit 77. At its left end in the figure screw 80 is carried in a sleeve 84 for rotation about the screw axis. At its other end, the screw is mounted to the drive shaft of the stepper motor, which in turn is carried on a sleeve 86. The dispenser device, worm screw, the two sleeves mounting

the worm screw, and the stepper motor used in moving the device in the "x" (horizontal) direction in the figure form what is referred to here collectively as a displacement assembly 86.

The displacement assembly is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an x axis in the figure. In one mode, the assembly functions to move the dispenser in x-axis increments having a selected distance in the range 5-25 μm. In another mode, the dispenser unit may be moved in precise x-axis increments of several microns or more, for positioning the dispenser at associated positions on adjacent supports, as will be described below.

The displacement assembly, in turn, is mounted for movement in the "y" (vertical) axis of the figure, for positioning the dispenser at a selected y axis position. The structure mounting the assembly includes a fixed rod 88 mounted rigidly between a pair of frame bars 90, 92, and a worm screw 94 mounted for rotation between a pair of frame bars 96, 98. The worm screw is driven (rotated) by a stepper motor 100 which operates under the control of unit 77. The motor is mounted on bar 96, as shown.

The structure just described, including worm screw 94 and motor 100, is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an y axis in the figure. As above, the structure functions in one mode to move the dispenser in y-axis increments having a selected distance in the range 5-250 μm, and in a second mode, to move the dispenser in precise y-axis increments of several microns (μm) or more, for positioning the dispenser at associated positions on adjacent supports.

The displacement assembly and structure for moving this assembly in the y axis are referred to herein collectively as positioning means for positioning the dispensing device at a selected array position with respect to a support.

A holder 102 in the apparatus functions to hold a plurality of supports, such as supports 104 on which the microarrays of regent regions are to be formed by the apparatus. The holder provides a number of recessed slots, such as slot 106, which receive the supports, and position them at precise selected positions with respect to the frame bars on which the dispenser moving means is mounted.

As noted above, the control unit in the device functions to actuate the two stepper motors and dispenser solenoid in a sequence designed for automated operation of the apparatus in forming a selected microarray of different-sequence polynucleotide regions on each of a plurality of supports.

The control unit is constructed, according to conventional microprocessor control principles, to provide appropriate signals to each of the solenoid and each of the stepper motors, in a given timed sequence and for appropriate signalling time. The construction of the unit, and the settings that are selected by the user to achieve a desired array pattern, will be understood from the following description of a typical apparatus operation.

Initially, one or more supports are placed in one or more slots in the holder. The dispenser is then moved to a position directly above a well (not shown) containing a solution of the first solution to be dispensed on the support(s). The dispenser solenoid is actuated now to lower the dispenser tip into this well, causing the capillary channel in the dispenser to fill.

Motors 82, 100 are now actuated to position the dispenser at a selected array position at the first of the supports. Solenoid actuation of the dispenser is then effective to dispense a selected-volume droplet of that solution at this location. As noted above, this operation is effective to dispense a selected volume preferably between 2 pl and 2 nl of the polynucleotide solution.

The dispenser is now moved to the corresponding position at an adjacent support and a similar volume of the solution is dispensed at this position. The process is repeated until the solution has been dispensed at this preselected corresponding position on each of the supports.

To dispense the next solution, the dispenser is positioned over a wash solution (not shown), and the dispenser tip is dipped in and out of this solution until the reagent solution has been substantially washed from the tip. Solution can be removed from the tip, after each dipping, by vacuum, compressed air spray, sponge, or the like.

The dispenser tip is now dipped in a second polynucleotide solution well, and the filled tip is moved to a second selected array position in the first support. The process of dispensing solution at each of the corresponding second-array positions is then carried as above. This process is repeated until an entire microarray of reagent solutions on each of the supports has been formed.

Fig. 5 shows a microarray device of the type useful in practicing the method of the invention. The device includes a glass substrate 138 having formed on its surface, a coating of a polycationic polymer, preferably a cationic polypeptide, such as polylysine or polyarginine. Formed on the polycationic coating is

a microarray 140 of distinct sequence polynucleotides, each localized at known selected array regions, such as regions 142.

The slide is coated by placing a uniform-thickness film of a polycationic polymer, e.g., poly-l-lysine, on the surface of a slide and drying the film to form a dried coating. The amount of polycationic polymer added is sufficient to form at least a monolayer of polymers on the glass surface. The polymer film is bound to the surface via electrostatic binding between negative silyl-OH groups on the surface and charged amine groups in the polymers. Poly-l-lysine coated glass slides may be obtained commercially, e.g., from Sigma Chemical Co. (St. Louis, MO).

To form the microarray, defined volumes of distinct sequence polynucleotides are deposited on the polymer-coated slide, as described in Section II. According to an important feature of the device, the deposited polynucleotides remain bound to the coated slide surface non-covalently when an aqueous DNA sample is applied to the substrate under conditions which allow hybridization of reporter-labeled polynucleotides in the sample to complementary-sequence (single-stranded) polynucleotides in the substrate array. The method is illustrated in the examples below.

To illustrate this feature, a substrate of the type just described, but having an array of same-sequence polynucleotides, was mixed with fluorescent-labeled complementary DNA under hybridization conditions. After washing to remove non-hybridized material, the substrate was examined by low-power fluorescence microscopy. The array can be visualized by the relatively uniform labeling pattern of the array regions.

In a preferred embodiment, each microarray contains at least $10^3$ distinct polynucleotide or polypeptide biopolymers per surface area of less than about 1 $cm^2$. In the embodiment shown in Fig. 5, the microarray contains 400 regions in an area of about 16 $mm^2$, or $2.5 \times 10^3$ regions/$cm^2$. Also in a preferred embodiment, the polynucleotides in the each microarray region are present in a defined amount between about 0.1 femtomoles and 100 nanomoles in the case of polynucleotides. As above, the ability to form high-density arrays of this type, where each region is formed of a well-defined amount of deposited material, can be achieved in accordance with the microarray-forming method described in Section II.

Also in a preferred embodiments, the polynucleotides have lengths of at least about 50 bp, *i.e.*, substantially longer than oligonucleotides which can be formed in high-density arrays by various in situ synthesis schemes.

III. Identifying Gene Expression Levels

This section describes a method which uses the polynucleotide arrays described above for identification of differential gene expression levels of multiple genes in test and control cells. Such differential gene expression can be determined, for example, between different tissues in an organism, or between samples from the same tissue or cell type in different states, such as activated/non-activated, infected/non-infected, cancerous/non-cancerous, affected/not affected, and diseased/healthy. Cells in these states are collectively referred to as "test" cells (e.g., activated, infected, cancerous, affected, diseased, etc.) and "control" cells (e.g., non-

activated, non-infected, non-cancerous, not affected, healthy, etc.). More generally, differential gene expression is determined using cells from a population of test individuals who have a particular condition or disease (test cells), and another population of control individuals who are "normal" for the condition selected, e.g., disease or infectious condition in the test individuals.

The method employs an array, and preferably a microarray of "target" gene sequences, or polynucleotides, which is then probed with reporter-labeled nucleic acids obtained or derived from the test and control cells. The amount of reporter-labeled nucleic acid that hybridized at each position in the microarray is determined. This signal reflects the amount or relative number of reporter-labeled nucleic acid fragments in the hybridization mixture that have a nucleotide sequence that is effective to hybridize with the nucleotide sequence of the immobilized target gene or fragment. Accordingly, if a particular messenger nucleic acid is represented at a relatively high level in the hybridization mixture, it will give a strong hybridization signal (high reporter level) at the position of an immobilized target having a sequence that is homologous to the complement of the reporter-tagged nucleic acid sequence.

The pattern of reporter signals may be quantitated and analyzed on a computer. As described below, the levels of gene expression, as determined by reporter levels bound to the genes in the array, are determined for a population of test or control individuals, to provide a statistical measure of genes in the array which are expressed at above- or below-normal values in test individuals.

A. Fabrication of Array

An array, such as a microarray, of immobilized target gene sequences is fabricated as described above. The DNA used in fashioning the array may be obtained from any of a variety of sources. Preferably, the array contains at least $10^3$ distinct gene sequences per $cm^2$ surface area. The array contains at least $10^2$ distinct gene sequences, and preferably $10^3$-$10^4$ distinct sequences.

The DNA at each location in the array has a unique nucleotide sequence, which may be known or unknown. For example, a genomic library can be constructed from DNA from which repetitive sequences have been self-hybridized-out. Alternatively, such a library may be purchased from commercial sources, such as Clontech (Palo Alto, CA). Clones from the library may be spotted on a substrate to fashion a set of arrays which includes all of the unique sequences represented in the library.

Clones from cDNA libraries may also be spotted onto a substrate to generate a microarray. Methods for the construction of cDNA libraries from a variety of tissues are well known in the art (Maniatis, *et al.*, Ausubel, *et al.*, Sambrook, *et al.*). Further, cDNA libraries may be conveniently obtained from several commercial sources, including Clontech (Palo Alto, CA) and Stratagene (La Jolla, CA).

Arrays may also be made with expressed sequence tags (ESTs) derived from various tissues (e.g., Orr). A set of arrays can be made for each tissue for which there exist ESTs, or a more general set of ESTs derived from two or more tissues can be prepared, depending on the application.

In a general embodiment of the invention, ESTs, or EST genes from multiple tissue sources are spotted onto a substrate to generate an array. Preferably, the array includes at least $10^3$ such EST genes. The EST genes in the plasmid vectors are amplified using primers directed against regions of the plasmid flanking the inserts. The amplified DNAs are purified with PCR clean-up kits (Quiagen, Chatsworth, CA), and spotted to generate an array as described above. An exemplary source of ESTs is the Merck-Washington University Consortium for ESTs (Merck and Co., Whitehouse, NJ), as well as through the Image Consortium of Lawrence Livermore Labs (Livermore, CA).

Each distinct gene sequence (e.g., single library clone or EST) is preferably disposed at a separate, defined position in the array and is present in a defined amount, typically between about 0.1 femtomole and 100 nanomoles. The spotting, or application of the DNA to the substrate, ~~is~~ may be done in duplicate to provide a measure for the consistency of the assay and validity of the results.

Following the spotting, the slides are typically rehydrated in a humid chamber for about 2 hours, snap dried on a hot plate at 100°C, rinsed to remove un-absorbed DNA, denatured, UV-crosslinked and treated with succinic anhydride as described in Example 1. The rehydration and snap drying are performed to facilitate the formation of a uniform, as opposed to an annular, distribution of the dried DNA. In another embodiment of the method, the DNA is spotted onto heated glass surface, which eliminates the rehydration and snap-drying steps and may achieve a similar uniform distribution of DNA.

B. Test Individuals Having a Shared Phenotype

Reporter-labeled copies of messenger nucleic acids, or fragments, are prepared from cDNAs or mRNAs obtained from "test" and "control" sources. As indicated above, the test sources are typically individuals (test individuals) having a shared phenotype that is not present in control individuals. Examples of such a phenotype include a pathogenic infection, such as a viral, bacterial or parasitic infection; a disease, such as a cancer, heart disease, diabetes, AIDS, an autoimmune disease, allergy, asthma, cardiovascular condition, various genetic diseases, degenerative diseases, and the like; and a predisposition to a disease or condition, such as a predisposition to heart disease, hypertension, diabetes, weight gain, stroke, neurodegenerative disease, psychiatric disease, and other diseases or conditions with an inheritable risk factor.

The shared phenotype may also include, for example, exposure to a known therapeutic agent or treatment modality, e.g., x-irradiation, or exposure to known environmental factors, e.g., a suspected factor in the workplace.

The control sources can be individuals (control individuals) who are "normal" for the test phenotype, i.e., individuals who are not affected with the disease or condition for which it is desired to construct a subarray, or any standard nucleic acid sample that provides a standard, reproducible hybridization signal at each array element.

C. Test Cells and Control Cells

The test and control cells used as a source for the reporter-labeled copies of messenger nucleic acids, or fragments, are typically those cells that are

directly affected by the disease or condition for which it is desired to generate a subarray of the present invention.

As examples, for use in constructing a subarray of genes whose gene expression levels are specifically related to a specific tumor condition, the test cells may be neoplastic cells, and the control cells, non-neoplastic cells of the same type.

For use in constructing a subarray of genes whose gene expression levels are specifically related to a genetic disease, the test cells may be cells from a tissue whose functioning is affected by the disease, and the control cells, cells from the same tissue in a normal individual.

For use in constructing a subarray of genes whose gene expression levels are specifically related to a virus-infected cell, the test cells may be virus-infected cells, and the control cells, uninfected cells of the same cell type from a non-infected individual.

For use in constructing a subarray of genes whose gene expression levels are specifically related to immune cells under immunological challenge, the test cells may be immunologically challenged immune cells, and the control cells, non-challenged immune cells of the same cell type.

For use in constructing a subarray of genes whose gene expression levels are specifically related to drug response in a given test cell type, the test and control cells may be the same cell type, in the presence and absence of the drug.

The invention also contemplates the use of cells other that those directly affected by the disease or condition as a source of test and control cells. Using the above liver tumor example, if it is suspected that certain genes in peripheral blood cells are upregulated

or downregulated as a result of a liver tumor, the test cells for generating a subarray of liver tumor regulated genes may be peripheral blood cells from an individual having a liver tumor, and the control cells peripheral blood cells from a normal individual, e.g., peripheral blood lymphocytes.

D. Preparation of Reporter-Labeled Nucleic Acid

Reporter-labeled copies of messenger nucleic acid fragments are prepared from cDNAs or mRNAs obtained from "test" and "control" sources. Messenger RNA, e.g., polyA RNA, may be reporter labeled by conventional methods, e.g., where the label is introduced at preferably the 5′ or 3′ end of the mRNA by suitable terminal transferase enzymes in the presence of fluorescent labeled nucleotides. cDNA may be reverse-transcribed from mRNA isolated from test and control cells, and may be used directly in reporter-labeled form, where the label is preferably introduced through the use of labeled nucleotides during reverse transcription and/or second-strand synthesis. Alternatively, the cDNA formed may be amplified by PCR or may be may be obtained from cDNA libraries that were generated from test and control cells. In cases where the reporter-labeled copies of messenger nucleic acid is obtained from a cDNA library, it is nevertheless considered to be "obtained" from the cell or tissue from which the library was made.

Messenger nucleic acids, or fragments, from the test and control cells may be labeled with different reporter moieties that can be independently detected with a minimum of signal cross-contamination. In one embodiment of the invention, the two sets of reporter-labeled messenger nucleic acids, corresponding to test and control cells, are hybridized simultaneously to an

array containing at least $10^2$ immobilized DNA fragments, each having a different nucleotide sequence. As is described more fully below, signals from the two reporter moieties may be detected either simultaneously or sequentially. Once normalized, the relative intensity of the signals reflects differences in the level of gene expression between the test and control cells.

Test and control cells suitable for the preparation of reporter-labeled copies of nucleic acid may be obtained directly from "test" and "control" individuals (e.g., as a tissue biopsy or blood sample), or they may be obtained from cell or tissue culture banks. Of course, cells that had been frozen or otherwise preserved may also be a suitable source of polynucleotides for the preparation of reporter-labeled messenger nucleic acid fragments.

Methods for the isolation of RNA and/or mRNA from cells or tissue, as well as for reverse transcription of mRNA into cDNA, are well known (e.g., Ausubel, *et al.*). In one embodiment of the invention, reporter-labeled nucleic acid is prepared during the reverse-transcription of mRNA into cDNA by substituting one of the four deoxynucleotides (A, C, T or G) with a reporter-labeled analogue of that deoxynucleotide. For example, the reaction can be carried out as described in Example 1, below, using fluorescein-12-dCTP (or lissamine-5-dCTP).

Reporter-labeled messenger nucleic acids may also be prepared by 5′ or 3′ end-labeling existing cDNA fragments (e.g., inserts of clones from a cDNA library) with a reporter moiety using known methods (e.g., Ausubel).

A number of different reporters may be employed, though the final detectable moiety on the reporter is preferably fluorescent. As discussed above, nucleotides having a fluorescent moiety attached are commercially available.

Fluorescence-based detection has several advantages over other types of reporter systems. For example, fluorescent signals do not scatter, permitting a closer spacing of the cDNA elements relative to spacing that could be achieved using radioactive or chemiluminescent detection methods. Furthermore, fluorescent signals can be multiplexed using different fluorophores for simultaneous detection of many hybridization reactions on the same array.

In certain cases, it may be desirable to use a fluorescently-labeled secondary reporter molecule that recognizes a primary molecule incorporated into or derivatized to the copies of messenger nucleic acid. A number of such systems are available. The most commonly used are biotin and digoxigenin. Either label can be easily incorporated into DNA probes and be detected using fluorochromes, which are available directly conjugated to anti-digoxigenin antibodies and to avidin. Kits for performing such labeling reactions are available, e.g., from Amersham (Arlington Heights, IL) and Boehringer-Mannheim (Indianapolis, IN).

The fluorescent detection moieties of the reporter system are preferably selected to have excitation maxima at a wavelength where the excitation light source can provide a strong excitation. For example, the system described in Example 1, below, employs a mixed gas multiline 10 W laser that generates spectral lines at a number of wavelengths including 488 nm and 568 nm. These wavelengths are near the excitation

maxima of the two fluorophores used (494 nm for fluorescein and 570 nm for lissamine.

As suggested above, fluorescence detection also allows for simultaneous determination of test and control expression levels on a single microarray. In this embodiment, test and control nucleic acids are labeled with different fluorescent reporters, and expression levels of each is measured independently on the array. The ability to independently monitor differently labeled nucleic acids on a microarray is illustrated in Fig. 6. The figure shows the hybridization pattern of the two yeast chromosome pools hybridized to an array of lambda clones with the inserts. A red signal in the figure indicates that the clone on the array surface originates from one of the largest six yeast chromosomes. A green signal indicates that the lambda clone originates from one of the smallest ten yeast chromosomes. Orange signals indicate repetitive sequences which cross hybridize to both chromosome pools. Control spots on the array confirm that the hybridization is specific and reproducible.

E. Hybridization of Reporter-Labeled Nucleic Acid to the Array

Hybridization of reporter-labeled nucleic acid to the array of immobilized target gene sequences is carried out using standard methods (e.g., as described in Maniatis, *et al.*, Sambrook, *et al.*, Ausubel, *et al.*) taking into account any special circumstances of a particular hybridization reaction. Typical hybridization conditions are given in Example 1. Such hybridization conditions are effective to hybridize the reporter-labeled nucleic acid to complementary-sequence target genes immobilized on the array.

Special circumstances relating to hybridization reactions employed in the practice of the invention include, for example, cases where the array is a microarray (e.g., a 5 $cm^2$ or smaller array) and benefits from small hybridization volumes (e.g., about 10 $\mu$l). Such a microarray spotted on, e.g., a glass slide, can be hybridized by placing the hybridization solution under a coverslip, and performing the hybridization in a humidified chamber, such as described in Example 1.

It will be appreciated that the level of reporter detected at any array position is directly related to the relative concentration of a particular-sequence reporter-labeled nucleic acid in the total mixture of reporter labeled copies of nucleic acids, i.e., the relative number of molecules of that nucleic acid species. Since the level of gene expression is related to the number of mRNA copies obtained from the test and control cells, and therefore to the number of cDNA copies produced from the mRNA species, the relative reporter level at each position on the array is directly related to the level of gene expression of that sequence in the test or control cells.

One important feature of the present invention, where the method is carried out using a gene microarray, is that volume of cell-derived cDNA material applied to array can be as little as 1-10 $\mu$l. For example, if a hybridization volume of 10 $\mu$l is used with a total of about 10 $\mu$g reporter-labeled cDNA, transcripts as rare as 1 part in 50,000 are detectable with fluorescent detection, without the requirement of amplifying the nucleic acid copies prior to hybridization. This ensures that the reporter levels measured at each position on the array reflect actual mRNA concentrations without severe distortions that may

be introduced by PCR amplification or using cloned cDNA species as a source of reporter-labeled nucleic acids.

Experiments performed in support of the present invention have demonstrated good quantitation can be achieved on a microarray using 2 $\mu$l hybridization volumes and 2 $\mu$g reporter-labeled cDNAs.

The specificity, lack of cross hybridization, and ability to detect expression at a 1:50,000 level in the method is illustrated in Figs. 7A and 7B, which show the fluorescent scans of the same array in a pseudocolor scale that was calibrated to reflect expression levels. Calibration of expression was performed using human acetylcholine receptor mRNA added to the *Arabidopsis* total poly-A mRNA prior to reverse transcription at a weight dilution ratio of 1:10,000 (arrays elements 1,2). Other *Arabidopsis* genes on the array whose expression levels were known from Northern analysis were also used for calibration purposes.

Fig. 7A is a scan of the array at a high photomultiplier tube setting for detection of rare transcripts down to a molar dilution ratio of as low as 1:100,000. It will be appreciated that adjacent dots (representing duplicate spots of the same cDNA clone) typically appear nearly identical in intensity, confirming the reproducibility of the spotting and immobilization procedures. Despite the high sensitivity setting used in this image, the negative controls (Yeast TRP4 at 95, 96; rat glucocorticoid receptor gene at 35, 36) did not produce a detectable signal. Note that the human acetylcholine receptor gene (1,2) produced a clear signal at a w/w dilution of 1:10,000.

Fig. 7B is a scan of the same array at a lower photomultiplier tube setting (intermediate sensitivity; detection of one in 10,000) to unsaturate the signals

of the more highly expressed genes on the array. This setting allows a linear detection of more abundant transcripts. Quantitation of both scans revealed a detectable range of expression levels spanning three orders of magnitude for the 45 genes tested. Details of the method are given in Example 1.

F. Detecting Hybridized Sequences

Following hybridization and washes at selected stringencies, the pattern of reporter levels for nucleic acids from the test cells is compared with that of nucleic acids from the control cells. The method used for detection of reporter levels, of course, depends on the reporter employed. Radioisotope reporters may be detected using, for example, autoradiography film or a "PHOSPHORIMAGER" (Molecular Dynamics, Sunnyvale CA). A "PHOSPHORIMAGER" can generate images similar to those shown, for example, in Figs. 6A and 6B. Similarly, chemiluminescent reporters can be detected using photographic film, while colorimetric reporters can be detected by eye and documented using a standard film camera or optical document scanner. The developed film images can be scanned into a microcomputer using commercially-available equipment and software, and analyzed.

The detection methods mentioned in the above paragraph typically have significantly lower spatial resolution than can be achieved using a fluorescence-based system. This limitation is not necessarily serious when the array being scanned is relatively large (e.g., a standard format 96-well or 384-well plate). However, when the array being analyzed is a microarray, it is preferable to use a fluorescence-based system.

An exemplary fluorescence-based system is described with respect to Example 2. The system uses a mixed gas laser as a light source, and has a computer-controlled X-Y stage for scanning the array in a raster fashion over a microscope objective. Standard excitation and emission ~~filter~~ filters are employed to analyze signals from different fluorophores. The fluorescence signals are collected using photomultiplier tubes, whose output can be conveniently digitized using an analog-to-digital (A/D) converter board in a microcomputer.

Fluorescence scanners such as is described are also available commercially. The scanner used in the experiments reported below is similar to the "Gene Chip Scanner" made by Affymetrix (Santa Clara, CA).

G. Analysis of Reporter Levels

The data obtained with the reporter detection device is analyzed to compare the pattern of reporter levels for nucleic acids from the test cells with that of nucleic acids from the control cells. Preferably, the fluorescence data are obtained in numerical form, such as from a photomultiplier tube connected to an A/D converter (e.g., as described in Example 1).

The data are typically stored, processed and/or further analyzed on a microcomputer. Typical processing includes displaying the data in a graphical form representing an image ~~that~~ of the entire array, with signal amplitude represented by the brightness and/or color at the corresponding location in the array. Figs. 7A, 7B, 8A, 8B, and 9A and 9B all show such "pseudocolor" images of the relative signal intensity at specific ~~location~~ locations in a microarray.

Figs. 8A and 8B show the detectable differences in gene expression between ~~wildtype~~ wild-type tissue labeled with

one fluorophore, and transgenic tissue labeled with a second fluorophore, determined at an intermediate-sensitivity (1:10,000) fluorescein scan of the cDNA array corresponding to the hybridization pattern of the wild-type *Arabidopsis* total cDNA. No detectable signal was observed from array elements 49, 50, and 1,2, indicating a lack of HAT4 expression and acetylcholine gene hybridization (acetylcholine gene DNA was added to match the HAT4 expression level). The positive control (rat glucocorticoid receptor gene; elements 35,36) showed a positive hybridization signal, while the negative control (the yeast TRP4 gene; elements 95,96) showed a lack of hybridization signal.

Figure 8B shows a matched-intensity lissamine scan of the same cDNA array corresponding to the hybridization pattern of the transgenic *Arabidopsis* total cDNA. Both HAT4 (elements 49,50) and the acetylcholine gene (elements 1,2; added to roughly match the HAT4 expression levels) showed strong hybridization signals. The positive control (the yeast TRP4 gene; elements 95,96) also had a strong signal, whereas the negative control (the glucocorticoid receptor gene; elements 35,36) had no detectable signal at this sensitivity setting.

Figs. 9A and 9B illustrate the ability of the method to determine complex gene patterns of expression in test and control cDNA mixtures. Fig. 9A shows a fluorescein scan of the cDNA array corresponding to the hybridization pattern of ~~the~~ root tissue cDNA. Fig. 9B shows a matched-intensity lissamine scan of the same cDNA array corresponding to the hybridization pattern of leaf tissue cDNA. The fluorescence intensities from the acetylcholine receptor gene (elements 1,2) on both arrays were matched by adjusting the photomultiplier tube settings.

The ratio of the fluorescence intensities of the two fluorophores in each spot provides a measure of the differential expression of that gene in leaf versus root tissues. Note the higher levels of expression of the photosynthesis-related chlorophyll binding protein in leaf versus root (elements 13, 14). Note also that other *Arabidopsis* genes were discovered to be more highly expressed in leaf than in root (e.g., element 91, 92) and, conversely, some genes were discovered to be more highly expressed in root than leaf (e.g., element 89, 90). Overall, 26 genes displayed greater than 5-fold differences in expression between control and test tissues.

The average or integrated value corresponding to the total signal from each region of the array is stored in a database for additional analyses. An exemplary additional analysis is the averaging of information obtained from a population of test individuals having a shared phenotype. Due to variations in the genetic make-up of unrelated individuals in a heterogeneous society, differences in the expression of a gene between any two individuals may or may not be significant. If such differences persist in a comparison of the averaged gene expression patterns from the two populations, it becomes more likely that the expression of that particular gene is related to the shared phenotype of the test individuals.

Further, it will be understood that the larger the number of individuals tested, the more significant the remaining differences in gene expression become. Standard statistical analyses may be applied to determine when the messenger nucleic acids from a sufficient number of individuals have been evaluated for differences in gene expression. Typically, samples

from at least 5, and preferably 20-50 different test individuals are assayed to obtain statistically meaningful data showing a significant elevation or reduction in reporter levels, when compared with control levels.

It will also be appreciated that the control expression levels for a particular array may be assayed with one population, and those control values can be used as a basis for comparison with a variety of test values, corresponding to different shared phenotypes, as long as no individuals in the control group exhibited any of the phenotypes for which that control is used as a reference.

In a preferred method of practicing the invention, microarrays for each of a number of test individuals are used to establish an "average" test pattern of gene expression levels for the genes on the microarray. Similarly, microarrays for each of a number of control individuals are used to establish an "average" control pattern of gene expression levels for the genes on the microarray. The test average pattern is then compared with the control average pattern, to identify those test genes which show significantly, typically at least 2 fold and up to 100 fold or more, increase or decrease in gene expression level with respect to control levels for the same gene.

Alternatively, average test and/or control levels of expression of genes on a microarray can be determined by combining equal amounts of reporter-labeled copies of messenger nucleic acids from each individual from a population of test or control individuals, and determining reporter levels associated with each gene on a single test array or a single control array (which may be the same as the test array).

The method just described takes advantage of the fact that even a single mutated gene or gene product, as well as a drug or other exogenous element, can have a large positive or negative effect on the expression levels of other genes in the organism, which can be detected and quantified using the methods detailed above.

No prior knowledge of the function of individual gene sequences is required in order to establish a correlation between a phenotype and gene expression pattern using the above methods. Rather, numerous hybridizations from individuals sharing a common phenotype are analyzed in a statistical fashion in order to establish a correlation. The correlation can be used for diagnostic purposes and/or monitoring of disease treatment, as will now be described.

IV. Subarray Device and Method

In another aspect, the invention includes a gene-array or subarray device and method of using the device to detect and/or monitor a disease condition.

The device includes a substrate and a subarray of genes which each show a statistically significant increase or a statistically significant decrease, typically at least 2 fold and up to 100 fold or more, in average gene level expression when compared with the average level of gene expression in a control cell type.

The genes in the device are those identified as showing a significant elevation or reduction in reporter levels in test cells, when compared with control-cell levels (Section III). Thus, for example, in constructing a gene-array device for detecting and treatment monitoring a given genetic disease, the method described in Section III is used to identify,

from a population of individuals with the genetic condition, those genes which show above- or below-average expression levels. These genes are then selected for use in the gene-array device.

The device itself may be constructed using the polynucleotide array-forming methods outlined in Section II. It is noted, however, that high density polynucleotide spotting is less important in this application, since many fewer genes will typically be required. Ideally, the number of genes whose expression levels ~~that~~ correlate with a particular gene condition or treatment method will be between about 5-50, although fewer or more genes may be involved.

Therefore the device, particularly in a microarray format, may include other gene regions, such as gene sequences that are not affected by the condition of interest, for use in establishing and normalizing to control levels of gene expression, or arrays designed for simultaneous detection of several different disease states.

In practicing the method, reporter-labeled copies of messenger nucleic acid are obtained from test cells associated with the physiological state or disease condition from an individual, as described above. The nucleic acid species from the test individual are then contacted with the gene-array device of the invention, whose genes are characterized by a statistically significant increase or decrease in gene ~~level~~ expression, when compared with the level of gene expression of the same gene in control cells. This contacting is carried out under conditions effective to hybridize the nucleic acid species to complementary-sequence genes in the array, similar to the hybridization conditions employed in the method detailed in Section III.

The levels of reporter associated with the genes in the subarray are quantitated, as above, allowing for the determination of a pattern of gene expression levels for the genes on the subarray device. In graphic representation, the pattern can take the form of a pattern of different colors, corresponding to different reporter levels and/or a pattern of different intensities of the same color. In digital form, the pattern may take the form of positive and negative digitized values, measured, for example, with respect to a zero-value control level. The control sample for the subarrays, if one is used, can be any reporter-labeled nucleic acid sequence that provides a standard, reproducible hybridization signal at each array element.

The test pattern is then compared to a diagnostic pattern generated preferably as the average pattern of a number of individuals known to have the diagnostic condition of interest. This average pattern can be constructed as above, either from a plurality of individual array patterns on the subarray device, or as the pattern of pooled nucleic acids samples from several test individuals.

Disease states can be identified, diagnosed and treatments monitored using such gene expression information. For example, patterns of gene expression triggered by mutations in oncogenes and tumor suppression genes can be used to characterize various cancers (Diamandis). Autoimmune diseases can, in part, be characterized according to the activation pattern of gene expression encoding the amplifying and proinflammatory cytokines (Osterland).

Further, subarrays of the present invention generated using the methods detailed herein may be used in drug development applications to measure

differential gene expression patterns of trial patient samples in response to drug candidates. Drug candidates that exhibit the desired effect on the genes of interest can be used as the basis of further drug design in a combinatorial or iterative drug development process. Once a drug is fully developed, the effectiveness of treatment for individual patients can be determined by hybridizing a patient's cDNA sample to ~~micro-arrays~~ microarrays containing the same genes used in the drug development process.

In one application, an array of cDNA clones representing genes is hybridized with total cDNA from an organism to monitor gene expression for research or diagnostic purposes. Labeling total cDNA from a normal cell with one color fluorophore and total cDNA from a diseased cell with another color fluorophore and simultaneously hybridizing the two cDNA samples to the same array of cDNA clones allows for differential gene expression to be measured as the ratio of the two fluorophore intensities. This two-color experiment can be used to monitor gene expression in different tissue types, disease states, response to drugs, or response to environmental factors. An example of this approach is illustrated in Examples 2, described with respect to Figs. 8A and 8B.

The following examples illustrate, but in no way are intended to limit, the present invention.

## Materials

### Buffers

**SSC (sodium chloride/sodium citrate), 20x**

3 M NaCl (175 g/liter)
0.3 M $Na_3$citrate-$2H_2O$ (88 g/liter)

pH adjusted to 7.0 with 1 M HCl

## Example 1

## Single-Color Fluorescence Detection of Gene Expression Patterns using Micro Arrays of Arabidopsis cDNA Clones

### A. Generation of Target DNA Fragments

Target messenger nucleic acid DNA fragments were made by amplifying the gene inserts from 45 different *Arabidopsis thaliana* cDNA clones and 3 control genes using the polymerase chain reaction (PCR; Mullis, *et al.*). The DNA fragments comprising the PCR product from each of the 48 reactions were purified using "QIAQUICK" PCR purification kits (Qiagen, Chatsworth, CA), eluted in $ddH_2O$, dried to completion in a vacuum centrifuge and resuspended in 15 $\mu$l of 3X sodium chloride/sodium citrate buffer (SSC). The capacity of the "QIAQUICK" purification kits is 10 $\mu$g of DNA; accordingly, each sample contained about 10 $\mu$g or less of DNA.

The samples were then deposited in individual wells of a 96 well storage plate with each sample split among two adjacent wells as a test of the reproducibility of the arraying and hybridization process.

### B. Fabrication of Microarray

The samples were spotted on poly-l-lysine-coated microscope slides (Sigma Chemical Co., St. Louis, MO) using the automated apparatus described above. The open-capillary printing tip loaded approximately 1 $\mu$l of each sample directly from the 96 well storage plates and deposited a 20 nl spot on each of 48 slides. The process was repeated for all 96 wells of the storage

plate with the spots on the each slide spaced about 500 μm apart.

After the spotting operation was complete, the slides were rehydrated in a humid chamber for 2 hours, snap dried on a hot plate at 100°C for 15 seconds, rinsed in 0.1% SDS to remove un-absorbed DNA, denatured in 90°C distilled water for 2 minutes and ultra-violet (UV)-crosslinked using a "STRATALINKER" (Stratagene, La Jolla, CA) set to a total (integrated) energy of 60 mJ. The samples were then treated with 0.1% succinic anhydride in a solution containing 50% N-methyl-pyrrolidinone and 50% 0.1 M Na borate buffer (pH 8.0) for 10 min to reduce non-specific adsorption of the labeled hybridization probe to the poly-l-lysine coated glass surface. The slides were rinsed in distilled water, air dried, and stored.

The positions of several specific elements in the 96-element array, and the reasons for their inclusion, are indicated in Table 1, below. The remaining elements of the array consist of known or unknown genes selected from an *Arabidopsis* cDNA library.

Table 1

| Element # | Name | Purpose |
|---|---|---|
| 1, 2 | Human acetylcholine receptor gene | Control for expression level |
| 13, 14 | Chlorophyll binding protein gene | Gene with known expression |
| 35, 36 | Rat glucocorticoid receptor gene | Positive and negative control |
| 49, 50 | HAT4 transcription factor gene | Gene with known expression |
| 95, 96 | Yeast TRP4 gene | Positive and negative control |

C. Preparation of Reporter-Labeled Messenger Nucleic Acid

The ability of the invention to monitor absolute, single-gene expression levels was investigated using a single-color fluorescently labeled nucleic acid sample hybridized to the *Arabidopsis* cDNA microarray fabricated as described above.

Total RNA was isolated from plant tissue of *Arabidopsis* using standard methods (Sambrook, *et al.*). PolyA+ mRNA was prepared from total RNA using "OLIGOTEX-DT" resin (Qiagen). Reverse transcription reactions were carried out using a "STRATASCRIPT" RT-PCR kit (Stratagene) modified as follows: 50 µl reactions contained 0.1 µg/µl *Arabidopsis* mRNA, 0.1 ng/µl human acetylcholine receptor mRNA, 0.05 µg/µl oligo-dT (21mer), 1X first strand buffer, 0.03 units/µl RNase block, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM fluorescein-12-dCTP (or lissamine-5-dCTP) and 0.03 units/µl "STRATASCRIPT" reverse transcriptase. Reactions were incubated for 60 ~~min~~ minutes at 37°C, precipitated with ethanol, and resuspended in 10 µl TE.

The samples were then heated for 3 ~~min~~ minutes at 94°C and chilled on ice. RNA was degraded by adding 0.25 µl 10N NaOH followed by a 10 ~~min~~ minutes incubation at 37°C. The samples were neutralized by adding 2.5 µl 1M Tris-HCl (pH 8.0) and 0.25 µl 10N HCl, and precipitated with ethanol. Pellets were washed with 70% ethanol, dried to completion in a "SPEEDVAC" (Savant, Farmingdale, NY) resuspended in 10 µl $H_2O$, and reduced to 3.0 µl in a ~~speedvac~~ SPEEDVAC. Fluorescent nucleotide analogs were purchase from DuPont NEN (Boston, MA).

D. Hybridization of Reporter-Labeled Nucleic Acid to Target DNA

Hybridization reactions contained 1.0 µl of fluorescent cDNA synthesis product (~2 µg) and 1.0 µl

of hybridization buffer (10X SSC, 0.2% sodium dodecyl sulfate; SDS). The 2.0 μl probe mixtures were aliquoted onto the microarray surface and covered with 12 mm round cover slips. Arrays were transferred to a waterproof slide chamber having a cavity just slightly larger than a microscope slide. The chamber was kept at 100% humidity internally by the addition of 2 microliters of water in a corner of the chamber. The chamber containing the arrays was incubated for 18 hr at 65°C.

The arrays were washed for 5 ~~min~~ minutes at room temperature (25°C) in low stringency wash buffer (1X SSC, 0.1% SDS), then for 10 ~~min~~ minutes at room temperature in high stringency wash buffer (0.1X SSC, 0.1% SDS). Arrays were scanned in 0.1X SSC using a fluorescence laser scanning device (see below).

E. Detection of Hybridized Sequences

The microscope used to detect the reporter-labeled hybridization complexes was outfitted with an Innova 70 mixed gas 10 W laser (Coherent Lasers, Santa Clara, CA) capable of generating a number of spectral lines, including lines at 488 nm and 568 nm. The excitation laser light was focused on the array using a 20X microscope objective (Nikon).

The slide containing the array was placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm × 1.8 cm array used in the present example was scanned with a resolution of 20 μm. Spatial resolutions up to a few micrometers are possible with appropriate optics.

In two separate scans, a mixed gas multiline laser excited the two fluorophores sequentially. Emitted light was split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu

Photonics, San Jose, CA) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes were used to filter the signals. The emission maxima of the fluorophores used were 517 nm (fluorescein) and 588 nm (lissamine). Each array was typically scanned twice -- one scan per fluorophore, using the appropriate filters at the laser source -- although the apparatus was capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans was typically calibrated using the signal intensity generated by an mRNA or cDNA control species added to the hybridization mix at a known concentration. For example, in the experiments described in Example 2, human acetylcholine receptor mRNA was added to the wild-type *Arabidopsis* poly-A total mRNA sample at a weight ratio of 1:10,000. A specific location on the array contained a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:10,000.

When messenger nucleic acids-derived probes containing two different fluorophores (e.g., representing test and control cells) are hybridized to a single array for the purpose of identifying genes that are differentially expressed, a similar calibration scheme may be employed to normalize the sensitivity of the photomultiplier tubes such that genes expressed at the same levels in the test and control samples display the same pseudocolor intensity. In one embodiment, this calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

It will be understood that where greater confidence in the absolute levels of expression is desired, multi-point calibrations may be performed.

F. Analysis of Patterns of Reporter Levels

The output of the photomultiplier tube was digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Norwood, MA) installed in an IBM-compatible PC computer. The digitized data were displayed as an image where the signal intensity was mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). Exemplary images generated using this method are shown in Figs. 7A, 7B, 8A, 8B, 9A and 9B.

The data were also analyzed quantitatively. In cases where two different fluorophores were used simultaneously, the data were first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid was superimposed over the fluorescence signal image such that the signal from each spot was centered in each element of the grid. The fluorescence signal within each element was then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for the above analyses was similar in functionality to "IMAGE-QUANT", available from Molecular Dynamics (Sunnyvale, CA).

## Example 2

### Two-Color Detection of Differential Gene Expression in Wild Type versus Transgenic *Arabidopsis* Tissue.

Differential gene expression was investigated using a simultaneous, two-color hybridization scheme, which served to minimize experimental variation inherent in comparing independent hybridizations. Two μg of wild-type *Arabidopsis* total cDNA that were labeled with fluorescein (as above) were combined with two micrograms of transgenic *Arabidopsis* total cDNA that were labeled by incorporating lissamine-5-dCTP (DuPont NEN) in the reverse transcription step and hybridized simultaneously to a microarray containing the same pattern of spotted cDNAs as described in Example 1.

To test whether overexpression of a single gene could be detected in a pool of total *Arabidopsis* mRNA, methods of the invention were used to analyze a transgenic line overexpressing the transcription factor *HAT4* (Schena, *et al.*). The transgenic *Arabidopsis* tissue was known to express HAT4 at levels of 0.5% of the total transcripts, while wild-type expression of HAT4 was only 0.01% of total transcripts (as previously determined by Northern analysis; Schena, *et al.*).

Human acetylcholine receptor mRNA was added to the wild-type *Arabidopsis* poly-A total mRNA sample at a weight ratio of 1:10,000 and into the transgenic *Arabidopsis* poly-A total mRNA sample at a weight ratio of 1:100 to roughly match the expected expression levels of HAT4.

As a cross-check of the negative controls, linear PCR (e.g., Cole, *et al.*, Manoni, *et al.*) was used to generate single-stranded fluorescein-labeled rat glucocorticoid receptor DNA and lissamine-labeled yeast TRP4 DNA. The two PCR products were added to the hybridization solution at a partial concentration of ~1:100. The two fluorophores were excited separately

in two separate scans in order to minimize optical crosstalk.

The array was then scanned separately for fluorescein and lissamine emission following independent excitation of the two fluorophores as described in Example 1, above. The results of the experiments are shown in Figs. 8A and 8B, discussed above.

## Example 3

### Two-Color Detection of Differential Gene Expression in Root versus Leaf Tissue.

In a similar experiment using the same labeling and hybridization procedures described above, 2 μg of total cDNA from *Arabidopsis* root tissue labeled with fluorescein were combined with two micrograms of total cDNA from *Arabidopsis* leaf tissue labeled with lissamine and were simultaneously hybridized to a microarray ~~micro array~~ containing the same pattern of target sequences described above. The acetylcholine gene mRNA was added to both poly-A total mRNA samples at 1:1,000 to allow for normalization of fluorescence intensities. The glucocorticoid and TRP4 controls were added to the hybridization probe as before. The results are shown in Figs. 9A and 9B, discussed above.

Although the invention has been described with respect to specific embodiments and methods, it will be clear that various changes and modification may be made without departing from the invention.